PACIFIC LINGUISTICS
Series C - No. 80

# STUDIES IN MICRONESIAN LINGUISTICS

Byron W. Bender
ed.


Department of Linguistics
Research School of Pacific Studies
THE AUSTRALIAN NATIONAL UNIVERSITY

Bender, B.W. editor. *Studies in Micronesian linguistics.*
**C-80**, xii + 527 pages. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.cover*

PACIFIC LINGUISTICS is issued through the Linguistic Circle of Canberra and consists of four series:

SERIES A - Occasional Papers
SERIES B - Monographs
SERIES C - Books
SERIES D - Special Publications



All correspondence concerning PACIFIC LINGUISTICS, including orders and subscriptions, should be addressed to:

The Secretary
PACIFIC LINGUISTICS
Department of Linguistics
Research School of Pacific Studies
The Australian National University
Canberra, A.C.T. 2601
Australia.



First Published 1984
Typeset by Mrs Judy Wise
Printed by A.N.U. Printing Service
Covers by Adriatic Bookbinders Pty. Ltd.
Maps drawn by Nigel Duffey, Cartography, Department of Human Geography, Research School of Pacific Studies, A.N.U.
The editors are indebted to the Australian National University for assistance in the production of this series.
This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN 0 85883 307 7

To

Bonnie Davis

1942-1981

whose zest for life

and

contributions to vernacular literature

enriched the lives of people

in Micronesia

and elsewhere

# TABLE OF CONTENTS

## THE CONTRIBUTORS

BYRON W. BENDER is a Professor of Linguistics at the University of Hawaii. In the fifties and early sixties, he lived in Micronesia and served as language adviser to the department of Education of the Trust Territory of the Pacific Islands.

JO-ANN FLORA is on the faculty of Trinity College in Washington, D.C. She worked on the Palauan language while she was a graduate student at the University of California at San Diego.

SHELDON P. HARRISON is on the faculty of the Department of Anthropology of The University of Western Australia. His research on the Kiribati (Gilbertese) language has been supported by the Australian Research Grants Committee. His research on Mokilese was done as a graduate student at the University of Hawaii.

ROBERT W. HSU is Assistant Professor of Linguistics at the University of Hawaii. He has done research on Palauan, Trukese, and Yapese.

FREDERICK H. JACKSON is on the faculty of Chiang Mai University in Thailand. He did original research on a number of the Trukic languages while he was a staff member of the Bilingual Education Programme for Micronesia 1975-82, and his dissertation is on the historical development of the Trukic group.

RODERICK A. JACOBS is Professor of Linguistics and ESL at the University of Hawaii. He has been doing research on the Kiribati (Gilbertese) language and several of the Trukic languages since coming to Hawaii in 1973.

JOHN THAYER JENSEN works for the Department of Education on Yap. He began working on the Yapese language as part of a Peace Corps training project when he entered the graduate programme in linguistics at the University of Hawaii in 1966.

LEWIS S. JOSEPHS was a member of the faculty of the Department of Linguistics of the University of Hawaii 1970-80. He is currently living and teaching in New York.

KEE-DONG LEE is on the faculty of Yonsei University in Seoul. He worked on the Kosraean language as a graduate student at the University of Hawaii.

ANN M. PETERS is Associate Professor of Linguistics at the University of Hawaii. Her research interests include language acquisition on the part of Micronesian children.

KENNETH L. REHG is Co-ordinator of the Bilingual Education Programme for Micronesia at the University of Hawaii. Before coming to Hawaii in 1970, he was Language Co-ordinator for the U.S. Peace Corps in Micronesia. His dissertation is on theoretical issues in Ponapean phonology.

HO-MIN SOHN is Professor of Korean at the University of Hawaii. He did research on the Ulithian language in the sixties for his doctoral dissertation, and on Woleaian in the early seventies for the Pacific Language Development Project at the University of Hawaii and the East-West Center.

ANTHONY F. TAWERILMANG is Assistant Director of Education for the State of Yap in the Federated States of Micronesia. He received the B.A. in Liberal Studies in 1974, and the M.A. in ESL in 1982, from the University of Hawaii.

JUDITH W. WANG is an instructor at the University of Hawaii where she is also completing her dissertation on the history of the Kosraean language. She was of considerable help in the editing of this volume.

# PREFACE

In the late sixties the editor of this volume reviewed in a survey article (Bender 1971) the linguistic dividends of the opening of Micronesia to the Western world following World War II: the literature beginning to accumulate at that time in English on the languages of the area. Four stages were noted:

(1) Beginning even before the American occupation of the islands preliminary studies were produced that were meant to be of value to new governers but were also of great interest to Western linguists, for whom the area was virtually unknown. Researchers of this era included Denzel Carr, Samuel H. Elbert, and Robert A. Hall.

(2) More systematic and in-depth work on some of the languages as part of the Coordinated Investigation of Micronesian Anthropology (CIMA) brought (inter alios) Paul L. Garvin and Saul H. Riesenberg to Ponape, Elbert to Kapingamarangi, Isidore Dyen and Ward H. Goodenough to Truk, and Arthur Capell to Palau and Sonsorol-Tobi.

(3) Civilian administration beginning in 1951 brought resident anthropologists, linguist-educators, and missionary-linguists, whose research efforts often had an applied side or took a back seat to the primary aims of their employment, but whose time in the islands often spanned years rather than months, and who were thus also able to contribute to the continuing task of describing the languages. Among these in addition to the editor of this volume were Charles-James N. Bailey, John L. Fischer, Eugene A. Nida, Edward M. Quackenbush, Alfred G. Smith, Donald M. Topping, and Gregory J. Trifonovitch.

(4) The decision to send hundreds of American Peace Corps volunteers to Micronesia in 1966 gave fresh impetus to basic linguistic research, and support from the United States government that continued, directly or indirectly, well beyond the appearance of the survey article, and through the additional stages which can now be identified:

(5) The language lessons produced for Peace Corps volunteers generally included grammatical notes and a glossary. The next step was to develop fuller reference grammars and bilingual dictionaries for use by the island governments in the production of documents and in the school systems. This effort began under a shoestring budget at the Pacific and Asian Linguistics Institute of the University of Hawaii, but soon gained the support of the Department of Education of the Trust Territory of the Pacific Islands. The East-West Center assisted in what came to be known as the Pacific Language Development Project by sponsoring speakers of Micronesian languages as research assistants in programmes that permitted them to work toward academic degrees at the University while serving as the chief resource people in the grammar and dictionary research. The first paper in this volume reports on the computer assistance developed for the

dictionary work by two University of Hawaii faculty members, and includes a list of the dictionaries that were written, most of which had a companion reference grammar project.

A prerequisite for the publication of dictionaries and reference grammars for use in island schools was the development of standard orthographies. This was accomplished by local committees that acted on the recommendations of the linguist teams. The paper by Jo-Ann Flora (who was not directly connected with the University of Hawaii activities) on schwa in Palauan is a sample of phonological analysis that was incorporated into the orthography recommendations made. The papers by Ho-min Sohn and Frederick H. Jackson on the Woleaian and Saipan Carolinian orthographies exemplify these recommendations and the varying extent to which they were followed in the committee decisions.

Four of the papers had their origins in a University of Hawaii graduate seminar on historical-comparative Micronesian linguistics conducted by Professor George W. Grace in 1971-72, while the dictionary and grammar research was still in progress and several years before those works appeared: those on the history of Ponapean phonology (Kenneth L. Rehg), and on Proto-Oceanic reflexes in Woleaian (Ho-min Sohn), Mokilese (Sheldon P. Harrison), and Kosraean (Kee-dong Lee and Judith W. Wang). Although Rehg's paper was revised for appearance in the *University of Hawaii Working Papers in Linguistics* in 1973, and Lee and Wang's has been revised in a qualified way for this volume, these papers do not reflect more recent attempts at reconstructing Proto-Micronesian or other subgroups of Proto-Oceanic such as are discussed under Stage 7 below.

The papers by Lewis S. Josephs on Palauan, and those by John Thayer Jensen on Yapese, Ho-min Sohn and Byron W. Bender on Ulithian, and Sheldon P. Harrison on segmental quantity in Mokilese are examples of spin-off from this stage addressed to other linguists. (The reference grammars were aimed at an educated lay reader.)

(6) The next stage that can now be recognised is one in which the products of the preceding stage are being interpreted to an audience of non-linguist island educators as part of broader training for programmes of bilingual education. An average of some 20 have participated in this project each year since its inception in 1975. It is estimated that by the end of 1983 some 50 participants will have been awarded University of Hawaii degrees.

The interaction between these participants and the faculty and advanced graduate students working closely with the project has made possible continuing research of a more reflective comparative or theoretical nature, of which the remaining papers in this volume are examples. Other examples are to be found in the proceedings of the symposia and international conferences on Austronesian linguistics that have been held since 1974 (see *Oceanic Linguistics* 12 and 13, March 1977, and Wurm and Carrington 1978). The termination of United States government funding which seems imminent as of this writing can be expected to diminish sharply the intensity of interaction that has characterised the Bilingual Education Project for Micronesia and may slow the research output. But papers of this sort can be expected to continue to appear from time to time. Attempts will be made to find support for continuing description, especially of Mortlockese and Nauruan, and for comparison and reconstruction of the sort to be found in Stage 7. Beyond that, no strong predictions will be hazarded, except to express the hope for a stage in which Micronesians themselves take the lead in writing about their languages and add psycholinguistic and sociolinguistic dimensions which have not been possible under conditions to date.

(7) The one other stage that can now be identified partially overlaps Stage 6, and dates from the appearance of Marck 1977, which gave impetus and focus to additional efforts at reconstructing a proto-language for those languages thought to be nuclear within Micronesia. A group of faculty and students at the University of Hawaii have been working cooperatively at this task, and as of this writing have compiled on computer file some 1400 tentative reconstructions. A few of these are cited in the paper on higher numerals by Harrison and Jackson, and Jackson's paper on Proto-Oceanic reflexes in Trukic gives a number of reconstructions for the Proto-Trukic subgroup. These papers also furnish a preview of the sort of results that can eventually be expected of these joint efforts, which could well be the basis for a sequel volume.

The organisation of this volume is genetic and geographic, proceeding (after the first four more general papers) from west to east, beginning with Palauan (a Western Austronesian 'outlier'), and proceeding to Yapese (whose status within Austronesian remains problematic). There follow in order papers from the five main branches of Micronesian: Trukic, Ponapeic, Kosraean, Marshallese, and Kiribati. Whether Nauruan should be considered a sixth is unresolved and beyond the scope of this volume. Other languages of Micronesia not treated in any direct way (in addition to Nauruan) include a second Western Austronesian 'outlier', Chamorro, and two Polynesian outliers, Kapingamarangi and Nukuoro.

## BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

MARCK, Jeffrey C.

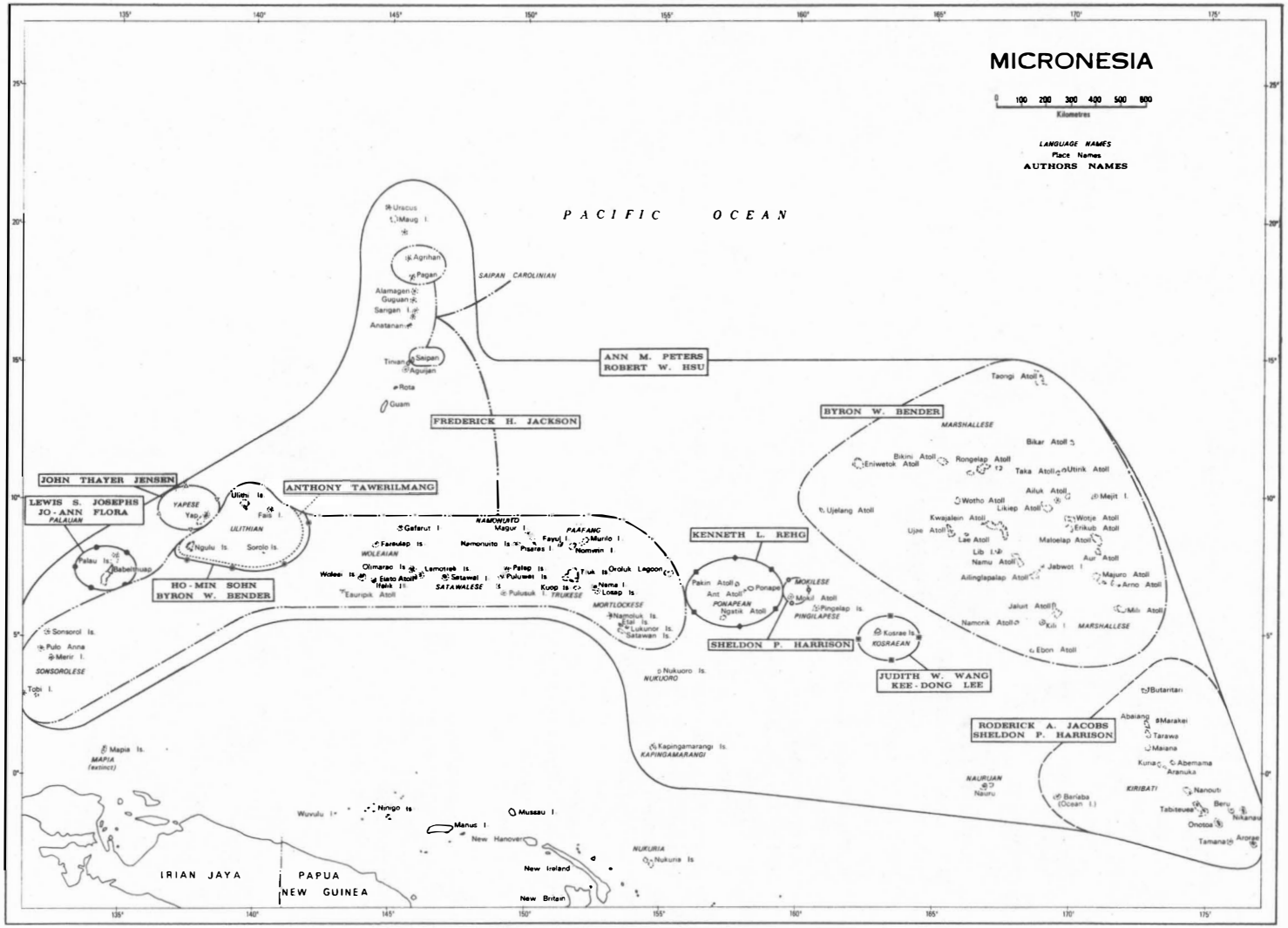
1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

WURM, S.A. and Lois CARRINGTON, eds

1978 *Second International Conference on Austronesian Linguistics: proceedings. Pacific Linguistics,* C-61.

# ABBREVIATIONS USED FOR LANGUAGE NAMES AND LANGUAGE GROUPS

| | |
|---|---|
| BON | Bontok (Philippine) |
| BUG | Bugotu (Santa Isabel, Solomon Islands) |
| CHA | Chamorro |
| CRL | Saipan Carolinian |
| Eng. | English |
| FIJ | Fijian |
| Ger. | German |
| HAW | Hawaiian (Polynesian) |
| ITW | Itawis (Philippine) |
| Jp. | Japanese |
| KAP | Kapingamarangi (Polynesian Outlier, Federated States of Micronesia) |
| KIR | Kiribati (Gilbertese) |
| KSR | Kosraean (Kusaiean) |
| MAO | Maori (Polynesian, New Zealand) |
| MC | Micronesian |
| MOK | Mokilese |
| MRO | Maranao (Philippine) |
| MRS | Marshallese |
| MRT | Mortlockese |
| MTU | Motu (Papua New Guinea) |
| NGK | Ngatikese |
| NGU | Nguna (New Hebrides) (Vanuatu) |
| NUK | Nukuoro (Polynesian Outlier, Federated States of Micronesia) |
| OC | Oceanic |
| PAL | Palauan |
| PAN | Proto-Austronesian |
| PEO | Proto-Eastern Oceanic |
| PLP | Pulapese |
| PMC | Proto-Micronesian |
| PN | Polynesian |
| PNG | Pingelapese |
| PNP | Ponapean |
| POC | Proto-Oceanic |
| PP | Ponapeic |
| PPN | Proto-Polynesian |
| PPP | Proto-Ponapeic |
| PTK | Proto-Trukic |
| PUA | Pulo Annian |
| PUL | Puluwat (Puluwatese) |
| REN | Rennellese (Polynesian Outlier, Solomon Islands) |
| ROV | Roviana (New Georgia, Solomon Islands) |
| SAM | Samoan (Polynesian) |
| SNS | Sonsorolese |
| Sp. | Spanish |
| STW | Satawalese |
| TBI | Tobi |
| TK | Trukic |
| TRK | Lagoon Trukese |
| TON | Tongan (Polynesian) |
| ULI | Ulithian |
| WOL | Woleaian |
| YAP | Yapese |




Location of language areas referred to by the authors

Bender, B.W. editor. *Studies in Micronesian linguistics.*
**C-80**, xii + 527 pages. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.cover*

# COMPUTERS AND MICRONESIAN DICTIONARIES: A CHRONICLE OF SYSTEMS-DESIGN FIELDWORK AMONG THE LEXICOGRAPHERS

Robert W. Hsu and Ann M. Peters

## 1. INTRODUCTION

The dictionaries of languages of Micronesia compiled in conjunction with research reported elsewhere in this volume were processed at the University of Hawaii with the help of specially written computer programmes. Though developed primarily to meet the needs of the Micronesian (MC) dictionaries, the programmes are of course indifferent to the source of the data, and have in fact also been used to process many dictionaries of other languages from around the Pacific. This article chronicles the development of these computer tools for lexicography. It is not, however, primarily a description of the tools but rather an account of their development. It tries to convey some of the flavour of this process by sketching in some background and then describing some of the problems that arose and the various attempts to solve them. It may be likened to a 'fieldwork history', of the sort often seen in the anthropological literature. Such accounts provide an additional dimension in which to evaluate and understand the 'formal' results of the work — in this case the computer tools developed and the dictionaries compiled. They are also frequently interesting in themselves because they expose an aspect of research that is not normally accessible to future investigators. The work reported here forms a chapter not only in the history of MC linguistics but also in the history of linguists' attempts to harness the computer to their complex and unformalised tasks, and should be of interest to other linguists and programmers contemplating similar projects. Another function of this article is to acknowledge the contributions to the development of computer-aided lexicography made by those lexicographers, many of whom are the authors of other papers in this volume, who have been the ultimate designers, as well as the guinea-pigs, victims, and beneficiaries of the programme systems whose evolution is described here. It has been their needs and ideas which have been responsible for this evolution, and which will influence the future course of this craft.

The system development work, as has already been hinted, took on many of the typical features of fieldwork, and since most linguists are more familiar with the latter than with computers it may be suggestive to view the work in the general terms of a linguistic fieldwork experience. The designers were the 'investigators'; the lexicographers' stated needs were the data from which generalisations were to be discovered and used to motivate the programmes and procedures, the 'grammatical description'. To the extent that these were usable by the lexicographers they embodied a proper analysis of the lexicographers' information

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 1-36.
*Pacific Linguistics*, C-80. 1984.

processing needs (their 'language'), whether stated or unconscious. Some of the early programmes and conventions proved definitely to be unsuitable, being reminiscent of the abortive but necessary initial attempts at description made by a linguist before the proper generalisations have been recognised. These descriptions are awkward to begin with, and are difficult to generalise or modify to accommodate the slightest additional input. Analogously, the early programmes were too specialised, and could not readily be generalised to answer needs slightly different from the one for which they were originally written. Other similarities with linguistic fieldwork are the following:

1. The lexicographers ('informants') were at first not consciously aware of all their needs, and those they were aware of were not necessarily formulated in data-processing terms ('theories').

2. The system designers were initially even more ignorant of these special needs, and were hesitant about imposing general external concepts on them — putting theory before the data.

3. The designers also frequently failed to appreciate the significance of every request of the lexicographer, just as a linguist does not always recognise the significance of every piece of information offered by the informant.

4. The development of the systems was influenced not only by the 'data' but also by the theoretical and practical tools available to the 'investigators', and their own biases and preferences.

5. Each party grew to better appreciate the outlook of the other. In fact, the association also led each party to better appreciate, and to refine, its own skills.

The following Section (2) will provide some background against which to view this 'fieldwork': 2.1. will describe the external circumstances of the work, 2.2. the technological tool-kit that was available, and 2.3. the background and interests of the 'investigators' themselves. These are three major, relatively independent, components of the background influences. In Section 3 we describe the 'fieldwork' itself, in terms of the interaction between the developing needs of the lexicographers (the unfolding 'data') and the design of the systems (the development of the 'grammatical description'). Elements of the background that have influenced this interaction are also mentioned. In an appendix are listed all the dictionaries that are mentioned in this narrative that have been published, together with relevant dates. Although this chronicle stops at 1976, the appendix has been updated to 1981.

## 2. BACKGROUND

### 2.1. Background and overview of the dictionary projects.

Over 25 dictionaries have been or are currently being processed by the computer at the University of Hawaii, of which about 14 are of languages of Micronesia. About 15 have so far been published, 10 by means of computer-controlled photocomposition (see 3.8.). They are the first extensive dictionaries to be compiled of these languages. There is little or no material written in these languages. The dictionaries are bilingual, with Eng. as the target language, and each is accompanied by an Eng. 'finderlist' (Eng.-to-source-language wordlist) generated by computer from the dictionary (for details see 3.3.). They are intended for general use by both Eng. speakers and speakers of the respective source languages, as well as by descriptive and comparative linguists and writers of teaching materials.

The completed dictionaries typically contain five or six thousand entries each, with a few having ten or twenty thousand. Each entry typically includes, besides Eng. definitions, such information as part-of-speech and other grammatical information, derivational privileges, loan source where appropriate, dialect and other variant forms, scientific identification of flora and fauna, and example sentences with translations. Some dictionaries further include synonyms and antonyms, morphological analysis of words where appropriate, idioms, derived and inflected forms where not readily predictable, reconstruction in proto-languages, phonemic transcription if the headword is given in some non-phonemic orthography, etc.

Each dictionary was typically compiled by one linguist in collaboration with several native speakers of the source language. A reference grammar of the language was usually produced as well. For several of the MC languages, work toward standardisation of orthography was concurrently undertaken, by special committees of native leaders and educators in consultation with the respective linguists. Data for each dictionary was gathered both in the field and, in most cases, also from native speakers brought to Hawaii. Often the linguist himself had at times to conduct the project from a distant university while data was entered or processed here in his absence. To date all computer-processing and data entry have been performed locally. Some of the dictionaries had been compiled over many years before they were committed to the computer here. Others were punched onto cards from the very beginning of data gathering.

All the dictionaries except for a block of seven from Micronesia belonged to separate unrelated projects and did not form a single coordinated lexicographical effort. These projects were aimed primarily at the compilation of dictionaries, and not at the development of theory or technique. In particular, they were not aimed at the development of computer techniques, though all of them assumed the availability of some computer support. Since existing computer programmes were initially not adequate to provide this support, special programmes were written by two linguist-programmers. These programmes generated a demand for greater support and attracted greater numbers of linguists to computerised lexicography. More programmes were written to satisfy their demands and the cycle repeated. Finally, after certain patterns of needs had become clearer and demand had risen to a critical level, a concerted effort was made to replace the heretofore scattered and specialised computer tools by a single coordinated system of programmes and conventions to satisfy a wide range of demands. Throughout this development, theory and technique of lexicography, at least as practised by the users of the programmes, have had to be refined and made more explicit. This has had an indirect unifying influence on the various projects.

The development of a unified programme system would not have taken place without the concentration of similar kinds of lexicographical activity at one place and time. For even though these dictionary projects seem, from the point of view of computer usage, quite different from those carried on elsewhere, they do, in retrospect, seem remarkably homogeneous among themselves. All of them have used the computer for storing and editing of dictionary entries, for printing multiple copies of dictionary files to take or send out to the field, for generating finderlists, alphabetising, helping with proofreading and consistency checking, and, in the case of most of the published dictionaries, for typesetting. They have not used the computer to process text in connection with discovering words or studying their usage and variation. Nor have they asked the computer for much help with the study of the information accumulated in the dictionaries themselves.

## 2.2. Technological background: hardware and software

The technological facilities available have played a key role in the directions the experiments have taken, much as the theories and techniques available to a linguist influence the work he does in the field. It will be helpful to be familiar with this background while reading about the 'fieldwork' in Section 3.

### COMPUTERS AND SOFTWARE

The computers in use at the University of Hawaii Computing Center when dictionary processing began in 1964-65 were the IBM7040 and 1401, with the latter used primarily for listing (printing out). Although we did obtain a version of the SNOBOL3 programming language and modified some of its inner workings so that it would run on the 7040, it ran too slowly, and hence was too expensive, to consider seriously for 'production' work (as opposed to small-scale experimental work). This left Fortran as the principal high-level programming language available for our use. The Sort/Merge programmes available on the 7040 system were used for alphabetisation.

In 1967 the Computing Center began a long association with the IBM 360 (and later the closely related 370) family of computers that continues to this day. The first model was a 360/50, acquired in January 1967. The early programmes written for the 7040 were transferred or rewritten for this machine. Fortran IV was still the language used, and an efficient Sort/Merge programme was again available on the new machine. The new programming language PL/1 had just become available, but, even though it was much more powerful than Fortran for our purposes, we at first avoided it for any serious work because the implementation was still unreliable and the documentation inadequate. During a later period however we programmed exclusively in that language. We also looked into COBOL and RPG as possible other languages to use but found them less suitable for our needs. Similarly we decided not to use LISP, which though extremely powerful for manipulating highly fluid data structures was poorly suited for string processing and inconvenient for formatting input and output.

In February 1969 the computer was upgraded to a model 65 (360/65) and in March we acquired a copy of the newly available SNOBOL4 compiler. We had been eagerly awaiting this language, with its convenient and powerful string manipulation facilities. We used it at first for some of the more complex formatting programmes but found it too slow and expensive on the computer for general processing. Nevertheless, partly because of the introduction of multiple task processing (MVT) on the 65, general computer costs were coming down rather significantly so that it became more economically justifiable to use the computer in ways that previously would have been prohibitively expensive. One of these ways was programming in SNOBOL4. The increase in complexity of the extraction conventions for finderlists (see 3.3.), in particular of the allowable environments of keywords and phrases, was a direct consequence of the availability of the more powerful features of SNOBOL4.

Another stimulus to programme development was the Computing Center's acquisition of the newly available SPITBOL compiler for SNOBOL4, in 1971. Identical programmes run from 6 to 10 times faster in SPITBOL than in SNOBOL4 and require much less memory space. In addition, after an initial trial period we found SPITBOL to be a well engineered and reliable system, in contrast to the early PL/1. Very soon we began writing exclusively in SPITBOL. It was now possible to develop, in less time, more complicated programmes that were at the

same time more economical to run, than would have been possible in any other language. It also became possible to respond to individual requests from linguists by quickly writing small, ad-hoc and inexpensive programmes. Furthermore, a few of the lexicographers themselves, after learning the rudiments of SPITBOL programming in a few lessons, were able to write simple inexpensive programmes for their own dictionary research.

The latest system of lexicographical programmes was written entirely in this language. That we undertook to design this system at all was due to the possibility, presented by SPITBOL, of writing the entire set of programmes in a reasonable amount of time. It probably would not otherwise have been attempted.

Most significantly, however, during a period of three or four years around 1971, the quality of service provided by the computer improved substantially even as costs were declining, due to increased hardware capability and reliability. Whereas previously we would often have had to wait for 12 hours (in effect overnight) for a job to be returned, the turn-around period steadily decreased until it was typically under an hour. The effect this had on our ability to experiment with new programmes and on the improvement of service to the lexicographers cannot be overestimated. Though usually taken for granted because it occurred gradually over a long period of time, the improved computer service has quite possibly been the single most decisive stimulus to the advancement of computer-lexicographical techniques, since it held forth the promise that, if armed with the appropriate programmes, the computer could be a dependable tool for daily use rather than a luxury to be resorted to only in case of extraordinary need. It thus opened the way for arming the computer with such programmes.

## INPUT-OUTPUT AND CHARACTER SETS

The keypunch has remained the most economical input device for us, taking into consideration not only cost but also convenience and the scale and sporadic nature of the dictionary processing. Perhaps the most constraining factor in the use of the keypunch for the linguist is the limited and inflexible character set. The first dictionaries were punched on the 026 keypunch, which had few symbols besides letters and numerals. After the 360 computer arrived, 029 keypunches began to be available, with the larger character set that is standard on the 360. We quickly assigned phonetic and typographical functions to many of the new characters. However there remained the inconvenient lack of upper and lower case distinctions on the keypunch. These were desired for readability, for distinguishing proper names, and for providing additional symbols for special uses. We investigated the possibility of an upper- and lower-case keypunch but the findings were not encouraging. There apparently existed such a machine, but it would not print on the card — neither lower case nor capital letters nor any other symbol. Since the printing was necessary for visual proofreading and locating of errors on the card, the non-printing keypunch was not considered acceptable for our purposes. However entering capital/lower case distinctions on the ordinary keypunch was not entirely ruled out even though only one case could be punched. We adopted the convention of punching the special character # before a letter that was to be interpreted as a capital. Letters without the prefixed # would be considered as lower case.

Lower case letters could not be directly encoded until a time-sharing system became available which could be accessed from a mixed-case terminal such as an IBM 2741 Selectric typewriter terminal. When this eventually came about,

it turned out for other reasons to be still impractical to use the terminal as an input device on a large scale (see below: TIMESHARING). Though we handled three or four dictionaries in this way, we were for most input purposes still effectively restricted to the keypunch.

What the 2741 terminal did make possible was the entry of data containing special characters that are available on Selectric type elements, e.g. a phonetic symbol element. A minor detail first had to be cleared up, however. There are two models of 2741 terminals, one with EBCDIC coding, the other with 'correspondence' coding. If one places a type element from an ordinary office Selectric on an EBCDIC-coded terminal, the letters typed out do not correspond to the letters on the keys depressed. On the correspondence-coded terminal such a type element prints the expected letters. Special linguistic type elements, such as Camwil's IPA symbol element, were of course designed for ordinary office Selectrics, and so must be used on correspondence- rather than EBCDIC-coded terminals. The latter however happen to be the standard model of 2741's available. After we unravelled this mystery and ordered the proper model of terminal, we were able to type in linguistic data with phonetic symbols and diacritics. The 2741 terminal thus opened up the possibility of entering material in certain languages with unusual character-set requirements, without turning to much more specialised and expensive equipment. Processing of dictionaries of several Amerindian languages of the Pacific North-west was thus made possible in this way, utilising a special element that had previously been designed for these languages by Laurence Thompson.

For printed output we have always relied on the Computing Center's high-speed line printers, which have increased in both speed and number (there are now 3) over the years. Any character that could be keypunched (except the exclamation and cent sign) could normally be printed on the line printer, but lower case letters could not. Thus upper and lower case material entered at the terminal, or keypunched text converted by programme to upper and lower case within the computer, could not be printed out in upper and lower case on the line printer. In 1972 the Computing Center rented a special print train (bank of type slugs), the so-called ALA train (designed by the American Library Association), which could be mounted on the printer when requested, and which contained lower as well as upper case letters. The train also contained certain diacritical marks, so that text with diacritics could be printed by appropriately overprinting the lines. The diacritics were still keypunched with special in-line characters, such as an apostrophe after a letter to represent an accent over it, and special programmes were written to cause them to print above the appropriate letters.

Material entered with a special type element, however, still could not be printed satisfactorily on the line printer because of the special characters required. Such printing still had to be done at the terminal itself, an extremely slow process. Other methods, such as purchasing all the necessary type slugs for the train or using a special multi-font terminal, would have been prohibitively costly for our scale of operations.

The ALA type font was however rather heavy in appearance and unpleasant to read, and furthermore, the train printed very slowly, a serious consideration when large dictionaries were to be printed and when the computer installation had only two printers. In late 1973 the Computing Center bought a TN print train, which had a much more pleasant upper and lower case type face and could print faster than the ALA train. We also managed to buy a slug with a few of the commonest diacritics to place in the train. We have since been using this train for upper and lower case print-outs.(see 3.5. for further discussion of character sets).

In summary, for data entry and printing out we have relied on the equipment standard at most computing installations, keypunches and line printers (augmented on occasion with an upper and lower case print train), and to a lesser extent on a typewriter terminal, because these machines have been available in-house or for a minimum commitment of funds. The need for special characters continues to be poorly met, at least for large volumes of output.

## TIME-SHARING

During 1970-71 the Computing Center installed several small interim or experimental time-sharing systems, with which we could do little more than experiment with on-line input and programme development. In 1971-72 the long-awaited general purpose TSO time-sharing system was installed. We were hoping to be able to process (e.g., edit) dictionaries on-line with this systen. Among other advantages of on-line editing would be the availability of immediate confirmation of every edit operation. Instant retrieval capabilities would also open up possibilities for using the dictionaries as an extensive research resource. Hardware problems in the early days, however, presented frustrating obstacles, resulting in little, if any, net saving in time and effort. There were difficulties with terminals, couplers, and telephone lines, as well as with the computer system itself. Over several years these problems have stabilised at a relatively low level, but still pose a significant deterrent to using the time-sharing system.

Another limitation to accessing dictionaries on-line has been the limited amount of on-line storage in the computing system. A typical large dictionary has had to be broken up into as many as 100 fragments, only one of which could be accessed on-line at any one time for retrieval and editing with the standard TSO facilities. Even though these limitations can now, with the steadily improving hardware situation at the Computing Center, probably be circumvented by appropriate programming, they and the unreliability of the system itself did succeed in steering us away from relying on time-sharing for dictionary processing. All the dictionary programmes in use now stay entirely within the 'batch' environment — that is, jobs are submitted, and later retrieved, at the Computing Center. Only a few dictionaries, as mentioned above, were or are being entered and edited at a terminal, at some inconvenience, in order to satisfy special character-set requirements.

The most practical use of time-sharing has turned out to be in the seemingly trivial area of programme maintenance. The many programmes in the system that has evolved are kept on disc files where they can easily be modified, augmented, and tested from a terminal. This is much faster, and more convenient and foolproof, than handling and transporting large decks of punched cards. This facility has been another significant factor in reducing the development time for the more recent programmes.

## PHOTOCOMPOSITION

Another technological development that has influenced our work is the growing availability of computer-driven photocomposition. In this technology, a properly prepared computer tape is fed to a specially programmed and equipped computer, usually at an installation specialising in this service, instructing it to project, by electronic, optical, and/or mechanical means, images of letters

and other symbols in various sizes, fonts, and positions onto special sensitive paper. The resulting camera-ready copy can be reproduced by standard methods for printing. Unlike manually composed galleys, computer-composed pages do not need to be proofread in detail, that is, if all the programmes involved have been properly tested.

We were anticipating the possibility of using this method of typesetting when the earliest dictionaries were first computerised. We corresponded with several firms as early as 1968, and visited the few small operations that were starting up locally. By the time one of the dictionaries was ready for the press in 1971, the cost of the commercial services had become low enough and the technical capabilities versatile enough for our purposes. Even then, we were embroiled in seemingly endless correspondence with the relatively small firm doing the work on the mainland, to iron out the misunderstandings, ambiguities, and mistakes that kept cropping up. The second dictionary was composed at a local firm, but the problems were no less simple to resolve since this firm had a much smaller machine. The difficulties with both firms were caused in part by the newness of the technology, and the consequent inexperience of the people involved, compounded by a lack of adequate documentation. The third and subsequent dictionaries to be photocomposed were again sent to the mainland, to a large and experienced firm which had more sophisticated equipment. This firm has been in much better control of the technology, so that we have had only relatively minor problems with these later typesetting jobs.

One effect of automatic photocomposition has been to lay on the lexicographers themselves the responsibility for the proper placement of typographical instructions in their own 'manuscript' or dictionary file. With more conventional methods of publication, this responsibility would have been borne by a special editor. Proofreading for typographical instructions in a printout must be taken particularly seriously since a small error can have more conspicuous effects in the printed book than a similarly small error in the text itself. The lexicographers have thus had to be very consciously aware of the intended typographical appearance of their material. This necessity has in turn been reflected in the development of additional computer tools, of the kinds described in 3.5. and 3.7., beyond those merely necessary to prepare the photocomposition tape, to ease the burden on the lexicographers (see 3.8. for further discussion of photocomposition).

## 2.3. Background of the system designers

Just as the skills and outlook of a linguist influence the course of fieldwork, so have the system designers' skills and outlook influenced the development of the systems of lexicographical programmes. Programmes for the first dictionary (Maranao (MRO)) were written by a succession of graduate students. Soon after, consulting and programming became the responsibility of two linguists who had had some previous experience with non-numerical computation, namely the present authors. There are now, in addition, student assistants to run, or help run, jobs for the lexicographers. The lexicographers themselves have in the main stayed away from actual programming and design. This section, then, concerns principally the skills and outlook of the authors of this paper and how these have influenced, and have been influenced by, computer processing of the Micronesian and other dictionaries.

Our outlook was dominated by an interest in promoting the wise use of computers in linguistics and in the humanities in general. Thus, even though we

felt that the field of lexicography could, in general, profit from use of the computer, we were cautious about recommending this route for every lexicographer. We insisted on trying to understand the particular motives of each prospective user, and on his trying to appreciate the advantages and disadvantages of processing his dictionary on the computer, since whether using the computer would result in any savings over using conventional methods depended on the nature of the project. For even if the computer can in some instances save some work, it can also produce additional work for the user, often in subtle and unanticipated ways. Several projects were, in fact, steered away from the computer as a result of such preliminary consultation.

Another aspect of our outlook that affected the evolution of the programmes was an interest in designing systems rather than in only writing programmes. As a general policy we tried to avoid responding to particular needs with special solutions. We were instead inclined to wait until a pattern of needs had become sufficiently clear before writing a general programme or devising a general convention to satisy the class of needs. A fieldworker, too, would not write a formalised description of each day's data, but would instead search in a larger corpus for patterns that could be described with greater generality. The lexicographers thus found themselves working not so much with programmers willing to help with every request as with independent investigators who had their own goals and interests to serve and their own ideas of what was interesting or worthwhile to design. The development of the finderlist programme, for instance, has been held back because we have been consistently unable to find any generalisations in the many requests for different features (see 3.3.). In contrast, the new data format (see end of 3.1.) is perhaps more general and complicated than necessary, though this has not proved to be a disadvantage.

Another motive for searching for generality was the realisation that, being linguists, we were in a position to attempt to understand the more general needs of our colleagues, and thus had a greater responsibility to develop general tools to meet these needs than to assist individual projects on an ad hoc basis. Even though we often performed the latter service in response to urgent individual requests, to have restricted our function to this would have been a waste of resources and opportunities. The effort to adhere to this general policy proved to be a profitable investment in the long run.

Though our training in linguistics was useful in this respect, not having ourselves compiled dictionaries was at times a handicap since we would often fail to grasp intuitively or to anticipate the difficulties of our clients or to evaluate the relative importance of their various needs. This resulted in much misunderstanding, backtracking and wasted effort on everyone's part. Furthermore, our 'elicitation' techniques in system-design fieldwork were poorly developed and we were generally unsuccessful in deliberately extending our 'corpus', as a linguist would normally do, in order to hasten the discovery of significant patterns.

The advantage of being relatively naive, on the other hand, was that we were less tempted to impose our own lexicographical technique and biases on our colleagues. The lack of preconceptions can be both an advantage and a hinderance to a fieldworker. We are now, of course, much better educated in the needs of a particular class of lexicographers. Nevertheless, unlike fieldworkers who learn to speak the language under investigation, we have still not become lexicographers ourselves, and have had no occasion to use any of our programmes. Thus there exists the possibility that their suitability is partly illusory and that more proper generalisations remain to be discovered. In sum, while being

linguists has often been an advantage, not being lexicographers has sometimes been a disadvantage.

Our computing skills and outlook also exerted a strong influence on the direction of the work. Our previous experience with computing was with small-scale modelling-type programmes, that is, with complexity rather than with quantity. We were familiar with some powerful programming languages, but not with assembly language or with file-management techniques. These were probably only minor handicaps but did influence our approach to the tasks, which were characterised, as it turned out, more by quantity of data than by complexity of processes. Thus we preferred to use higher-level string- and list-processing languages. We avoided assembly language and even felt that Fortran and PL/1 were unsuitable, though these were the only languages available to us at first. We were instead led to explore the possibilities of the SNOBOL series of languages, and eventually found SPITBOL to be the most suitable language for our use, where someone else might have remained with PL/1 or assembly language. We were, in fact, instrumental in getting SPITBOL installed at this computing centre. This choice, in turn, strongly influenced the development of our programmes. Our file structures and file management however remain of the most elementary sort available on the 370 computer system. As a consequence, when a dictionary tape is processed by our programmes, as when it is edited or printed out, it must be processed in its entirety, and sequentially, from beginning to end. For small dictionaries this constraint has not been intolerable, but for much larger ones, say those containing tens of thousands or hundreds of thousands of entries, our present mode of operation would be extremely cumbersome if not completely impractical.

We also were inclined to use available systems whenever possible. We checked the available tape editing programmes which come with most computer systems, but finding them too inflexible for our purposes we wrote our own (see 3.2.). However, for reasons mentioned above we hesitated to write our own editing system for time-sharing use even though the existing system was, for our purposes, severely restricted (see TIME-SHARING, above). We have instead resigned ourselves to living with the existing on-line editing system and, when necessary, juggling large numbers of small files. For sorting we could rely on the efficient Sort/Merge facility of the computer system. To permit alphabetical ordering different from that of the machine we developed a strategy for generating multi-level keys, which we called 'sort handles', preparatory to sorting. This method relied on familiar string-handling techniques, those for which the SNOBOL languages are admirably well suited, while still taking advantage of the efficiency of the Sort/Merge programmes. It did not require either meddling with the sort programmes or writing our own. It has in fact turned out to be a very versatile technique (see 3.6.).

Though we were in favour of using existing systems as much as possible, we stopped short of importing and installing, on our own, large systems developed elsewhere for file maintenance and retrieval, because of the difficulty of obtaining sufficient information to evaluate them and because of the suspicion, based on past experience, that by the time an imported system is successfully installed and learned one could have written one's own more directly applicable system. In retrospect this seems to have been the right decision.

We also had high hopes initially for time-sharing in dictionary processing, and actively sought to have a system installed here. When a general-purpose system was finally operational, however, various impediments, as noted above, prevented the realisation of its full potential for dictionary processing. In the meantime we have taken advantage of the system for less demanding, but nevertheless quite significant, uses.

## 3. DEVELOPMENT OF NEEDS AND OF PROGRAMME SYSTEMS

The data processing needs of the lexicographers, and the system designers' responses to them formed the substance of the 'fieldwork' experience. We have divided these needs into several classes and will discuss them in chronological order.

### 3.1. Information structure

Before information can be encoded (e.g. keypunched) for computer processing, conventions must be established for signalling the structure of the information. In punching a dictionary onto cards, for instance, conventions are needed to differentiate the entries one from another, and to signal the different parts of an entry (headword, definition, etc.). Developing such conventions, or data formats, for increasingly complex dictionary entries has been one of our central concerns, since the actual programmes must be designed around these conventions.

The first dictionary to be processed here had a very simple structure. It was the Maranao (MRO)-Eng. dictionary, punched in 1965. Each entry consisted of a headword and one or more Eng. words or phrases which served as definitions, with, in some instances, illustrative sentences. The conventions for punching were as follows: Each entry began on a new card, the headword starting in column 1 and followed on the same card by the definitions. Each definition was prefixed with an asterisk, which served to separate it from the previous definition or from the headword, as well as to mark the Eng. words to be extracted by a programme to make an Eng.-MRO finderlist. A typical card in this format would have looked like this:

```
KAKEP *METAL STRIP USED FOR BINDING*GO NEAR*SURROUND*ENCLOSURE*CLOSE
```

Illustrative sentences were kept in a separate file from the definitions. The headword was again punched on a card, but this time it was followed by the illustrative sentences, e.g.,

```
KAKEP/(1) KAKEPEN KA ON SO PINTOQ. CLOSE THE DOOR. (2) SO MANGA*
* KAKEP O BARUIL NA OADIA. THE BANDS OF THE BARREL ARE MADE OF*
* STEEL.
```

The stars at the ends and beginnings of cards were a means of indicating continuation of an item onto additional cards. This convention was also used for continuation of definitions. These two files were kept on two tapes in the same form in which they were punched, and they were sorted together before publication.

Keypunching of the Marshallese (MRS) and Chamorro (CHA) dictionaries began in 1966-67. Unlike the MRO dictionary project, these two projects had publication only as a very remote goal. Rather, the authors wished to have a computer-readable file for the lexical data that they had accumulated on slips over a number of years. The file was to be further revised and expanded indefinitely. They needed flexible conventions which would allow them to enter diverse types of information (e.g., etymology, variant forms, part of speech, definition, illustrative sentences, etc.) for each lexical entry and which would allow addition of new types of information in the future. These were therefore to be 'lexical files' rather than dictionaries. An actual dictionary for publication might, we envisioned, someday be extracted from the information accumulated in such a file. In the meantime it could continue to grow and could be used as a research and teaching resource.

For a number of reasons we decided to start with a punched card oriented system rather than with one that depended on the use of tape and hence the computer. The file would be kept on punched cards in boxes, corrections would be made by punching new cards and manually substituting them for the old cards, and new entries would be manually inserted in the proper alphabetical position in the boxes. The cards could be printed out on paper by the IBM 407 Accounting Machine — a noisy electro-mechanical monster that in those days still inhabited most computing centres. If the cards should be dropped or otherwise disordered, or if a batch of newly punched entries had to be alphabetised, they could be physically sorted on another electro-mechanical device, the model 83 sorter. A set of sorted new entries could be merged with the main file by using a third, now obsolete, machine, the model 87 collator.

This decision stemmed partly from a belief that being able to handle one's own data cards would restore some of the traditional flavour of being close to one's data that was lost through mechanical processing. Using a punched card file, the linguist could always have physical control of the data, could himself perform, manually and on the machines, the ordinary operations such as updating, sorting, and listing, and would only depend on a programmer and on the computer for more specialised processing. Since the volume of data and of activity on the two dictionaries was then rather low and the future course of the projects was not at all clear, such a humble technological beginning was, we felt, tolerable and appropriate. Only as definite limitations were felt would we seek or develop more powerful techniques. Furthermore, the computer in those days was not, as noted earlier, particularly accessible, and would not have saved a small project very much time.

As a consequence of this decision, the card format had to satisfy a number of conditions besides being easy to punch in:

1) A simple listing (print-out), made on the 407, of the cards as punched should be relatively readable and usable — that is, no formatting programme should be required for just getting a print-out.

2) Each card should carry on it all the information necessary to determine uniquely its proper place in the file. This information would ensure that cards removed from the file or misfiled could be properly replaced. It would also permit mechanical sorting of the cards.

3) The format should be flexible enough to accommodate varied types of information and varying amounts of information in each entry.

(The MRO format had not been designed to satisfy any of these requirements.)

Requirement 2 could be satisfied by punching a unique serial number on each card, but this would have made it next to impossible to expand or insert entries. Giving each entry a serial number would be similarly constraining. But since the headword of an entry had to be punched anyway, and the entries were to be kept in alphabetical order by headword, we decided to use the headword, duplicated into a special field on every card of an entry, as the main sequencing field of the file. A fixed-length field of 18 columns (columns 3 through 20 on the card) was assigned for this purpose. There was a provision for longer headwords but only their first 18 characters would be significant. When two headwords were spelled exactly alike, we required them to be disambiguated by suffixation of numerals following a dollar sign, e.g., $1, $2, etc. We called these 'subscripts' in anticipation of their typographic realisation in a future book. An additional field of two columns (columns 1 and 2), of which usually only the first was punched, was used to subsequence the cards within an entry. This field and the headword field together satisfied requirement 2.

Each type of information was punched on a separate card or group of cards in columns 21 through 80, and the contents of the subsequencing field, usually a single letter, assumed the additional duty of identifying the type of information on the card. For each dictionary the linguist assigned a letter or block of letters from the alphabet to each type of information that might be entered into the dictionary. Thus B might signal grammatical information; E etymological information; G through L definitions, which might need several cards; similarly Q through U for illustration sentences; and so on. These letters or blocks of letters, assigned to signal different types of information, were known as 'bands' because they corresponded to horizontal strips in the print-out of an entry (see example below). (This term contrasted with 'fields', which are vertical, or columnar, divisions such as the two-column subsequence field or the headword field.) Each information type then also came to be known as a 'band', as did the actual contents of such a strip in a particular entry. Each dictionary compiler was free to divide up the alphabet and assign bands in whatever way he chose, perhaps leaving certain letters unassigned in anticipation of possible future addition of new types of information. In practice it was not hard to become accustomed to what the different letters stood for in one's own dictionary. Not every band, of course, needed to be present in every entry: if there were no illustrative sentences in an entry there would be no cards punched for that band. Adding or changing material in a band that was already punched involved punching entire new cards to replace one or more of the old ones. When this resulted in having to insert additional cards between existing ones in a band, the additional cards could be punched with a 1, 2, etc. in the second column of the subsequence field: M, M1, M2, etc. all came before N. The two-column subsequence field thus served both to sequence the cards within an entry and to encode the type of information that was on the card. That is, it helped to satisfy requirement 2 and at the same time provided the flexibility of requirement 3.

Two MRS entries might have looked like this, as listed on the 407:

```
B DA"PDE'P          DRABDREB
H DA"PDE'P          1(-I), 2(INF, TR DA"PIJ), 3, 5(DA"PDIPIJ), 7, 8,
H1DA"PDE'P          10(DA"PIJEK)
N DA"PDE'P          *HOLD; *GRIP; *RESIST; *CLASP; *CONTROL; *KEEP;
O DA"PDE'P          *LINGER; *MAINTAIN; *NEGATIVE; *RETAIN; *ADHERE
R DA"PDE'P          #KATE EOK DA"PDE'P BWE WA IN EJ BUUL'.  'DO YOUR
S DA"PDE'P           BEST TO HOLD ON BECAUSE THIS VEHICLE IS GOING F
T DA"PDE'P          AST.'
B DDE'B             DREDRO"B (RAT.) EDDRO"B (RAL.)
D DDE'B             W: YE'DDE'B, E: DE'DE'B
H DDE'B             1(-I), 2(INF, TR DIBO"J), 3, 4, 5(DIBDIBO"J), 7,
H1DDE'B             8
N DDE'B             *HUSK *COCONUTS; *PIERCE WITH *HUSKING STICK OR
O DDE'B             *SPEAR; *JABBING
```

The two headwords in this example are DA"PDE'P and DDE'B. The bands are as follows: B, old MRS spelling of the word; D, dialectal variants (only the second entry in this example has this band); H, derivational information; N through Q, definitions; R through T, illustrative sentences.

Cards were easy to punch in this format and the listing was reasonably readable, i.e., requirement 1 was at least not grossly violated, especially if the print-out was double-spaced. Nevertheless we soon did write a simple print-out programme that could optionally be used, to perform a few formatting functions such as skipping a line between entries, suppressing the repeated headword field, numbering pages, etc.

Afterwards, when these dictionaries grew so large that they had to be committed to tape, this card-oriented format was retained. Each 'record' on tape corresponded to one 80-column card. Such a tape file is known as a 'card-image' file. Even though on tape there was no danger of cards, or records, being dropped, the sortability criterion was still relevant: requirement 2 is just as necessary for sorting by computer as for sorting on the 83 machine. The format was retained also so that the finderlist and other programmes that had by then been written would not have to be rewritten, but could be used for both card and tape files.

The Nukuoro (NUK) dictionary was keypunched beginning in 1966. Due to its size it was immediately transferred to tape rather than kept on cards. It had different information structure requirements from the MRS and CHA dictionaries. While the number of different bands (types of information) was small and fixed, the entire file had to be sortable not only by headword but also by root. Each headword was affiliated with one or more roots through different derivational relationships. Thus, while requirement 3 could be relaxed, requirement 2 was compounded.

The purposes of putting this dictionary on the computer were also different from those for computerising the previous two dictionaries. The NUK dictionary was already approaching its projected size before it was keypunched. It was now proposed to refine some of the definitions and to complete some of the derivational sets with the aid of computer print-outs and an informant, as well as to generate a finderlist. The prospect for indefinite growth that characterised planning for the previous two dictionaries was absent.

Although the NUK entries could have been distorted to fit the band format and then been reconstructed, by special programmes, for sorting and printing, we designed a new format in order to minimise such programming. The format was however convenient neither for keypunching nor for editing; it was convenient for sorting. Data on tape in this format could be sorted directly, in either of the required orders. The following example shows a few entries in an approximation of this format:

```
GANO/A      B   GANNUI          A   1 THE GELATINOUS LAYER IN
GANO/A      B   GANNUI          A   2 THE KERNEL OF THE IMMATURE
GANO/A      B   GANNUI          A   3 COCONUT
NUI$1       XCO GANNUI          A   1 THE GELATINOUS LAYER IN
NUI$1       XCO GANNUI          A   2 THE KERNEL OF THE IMMATURE
NUI$1       XCO GANNUI          A   3 COCONUT
GAO         R   GAOGAO          NL  1 SIDE OF; THE PART SURROUNDI
GAO         R   GAOGAO          NL  2 NG; ENVIRONS OF; SIDE OF
GAO         R   GAOGAO          NL  3 THE ABDOMEN
```

The fields are, from left to right, 1) root, 2) derivation type code (which gives the morphological relationship between the headword and the root), 3) headword, 4) word class code, 5) sequence number, 6) definition. The sequence numbers were necessary to keep the continuation cards in order during sorting. This example shows that the headword GANNUI has affiliations to two roots, GANO/A and NUI$1, in the relationships B and XCO respectively. Its word class, A, and definition, contained on three cards due to the shortness of the definition field, are replicated, one set for each root.

This excerpt is in order by headword (sorted by fields 3, 1, and 5). If the file were reordered by the roots (sorted by fields 1, 2, 3, and 5), these two sets of records for GANNUI would appear in two separate places in the printout because the roots GANO/A and NUI$1 would have been sorted apart. A third poss-

ible way to sort this file would be by word class code (sorting on fields 4, 3, 5), giving a list classified by word class.

After sorting in any of these ways the file was formatted for printing by simple programmes (requirement 1, direct readability, is not well satisfied by the above format). In return for the convenience of direct sortability of the file in all these ways, punching of the original and correction cards was made more difficult. This was due to the rigid data format, the short definition field requiring frequent continuation cards, and the consequent necessity of entering multiple copies of data or corrections wherever the same information was duplicated (and in some cases repeated three or four times). However we thought that this was an acceptable inconvenience for a dictionary that was going to be entered only once from beginning to end with probably very few corrections to be made afterwards.

In actual fact, the editing of the dictionary continued over several years, spurred on by computer-aided checking of the kinds described in 3.7. The format had to be drastically revised twice, and the programmes rewritten (in PL/1), to facilitate this process. Even so, work on the dictionary was in certain respects curtailed or discouraged because of time and expense. In the revised format redundant information was not stored on tape — requirement 2 was abandoned. Each field (headword, root, etc.) of an entry was stored only once. Instead, a special set of subprogrammes was written to convert between the different formats used for tape storage, card punching, sorting, processing within the computer, and printing. Thus by calling on the appropriate subprogrammes a tape for sorting satisfying requirement 2 could be reconstructed any time it was necessary to do any sorting, and entries to be edited could be converted into the processing format as they were read into the computer from tape, etc.

Later the similarly structured but smaller Kapingamarangi (KAP) dictionary was cast in the same scheme and processed in about two years' time (from 1971 to 1973). Since then this data format has not been used because no dictionary of a language so derivationally rich as these two Polynesian outliers has since been processed.

Meanwhile, the Bontok (BON) dictionary, already collected in two large typescript volumes, was being keypunched, beginning in 1970. The purpose of computerising this dictionary seemed to be similar to that of computerising the NUK dictionary: a final stage of expansion with the help of an informant, finderlist generation, and editing for publication. It too was committed, due to its initial size, to tape right from the beginning. But the structural requirements seemed at first to be pretty well met by the old band format. A few entries did, however, present a minor problem. When a word could be used both as a noun and as a verb for instance, the entry for the word would have to contain two sets of bands, each set containing, say, a part-of-speech band, a definition band, and an illustrative sentence band. To accommodate the two sets of bands by repeating the band letters in the same entry (e.g. A G L P A G L P) would have destroyed the useful sequencing property, within an entry, of the band letters (part of requirement 2). On the other hand, renaming the second set with letters coming later in the alphabet (e.g. Q S U Y) would have degraded the mnemonic value of the letters and would have required modifications in some of the programmes. Furthermore, the alphabet would have been hard pressed to accommodate a third or fourth set of band letters. It was evident that in order to stay within the old framework we would have to make two separate entries for the same headword. But they would have to be kept apart, during sorting and other processing, in some way. We adopted the convention that in the first entry the headword would be suffixed with '(A' and in the second '(B'. When

the file was sorted these suffixes kept the two 'entries' separated but next to each other in the proper order, and in the printout they signalled to the reader that the 'entries' belonged together. These 'entries' correspond to what were later to be known as different 'modes' of a single entry (see end of 3.1., the Palauan (PAL) example).

Another structural requirement that surfaced later in the BON dictionary was not so easily resolved within the old framework. This was the problem of handling subentries. These are entries which do not appear in their proper alphabetical order in the file but rather under the entry for the main word from which they are derived. We could find no way to punch subentries in the old format without abandoning requirement 2, that the file as keypunched should be sortable. At that time this requirement was also being assailed by the emerging need to sort dictionaries in orders other than the machine alphabetical order (see 3.6.). With this background, the problem of proper ordering of subentries was seen as part of the general need for 'unusual' ordering of entries, for which we already had a strategy, rather than as one requiring expansion of the data format. Abandoning requirement 2, then, we adopted the following convention. A subentry would be punched in the same way as a normal entry, except that it would have a special 'root' band, labelled by '$' instead of a letter, and in this band would be entered the root word under which the subentry was to appear. With this convention a subentry for the word PATAY, a derivative of ETEY, might have looked like this:

```
$ PATEY          ETEY
G PATEY          +V
M PATEY          TO BE EXTREMELY PAINFUL
```

To prepare a file for sorting, a programme would find every subentry, which was marked by having the $ band, and place a copy of the root followed by a copy of the headword of the subentry (e.g. ETEY PATEY) in a new field, called the 'sort handle'. This field was added to the 80-character card image records on tape. (Records on tape are not restricted to 80 characters in length.) If an entry was not a subentry, the headword alone would be copied into this field. Thus the entry for the root word ETEY and the subentries for the derived words P'ATAY and PATEY would be assigned sort handles as follows:

| headword | handle |
|---|---|
| ETEY | ETEY |
| P'ATAY | ETEY P'ATAY |
| PATEY | ETEY PATEY |

Sorting by the handle field instead of the headword field caused the subentries to follow immediately the root word ETEY rather than to come somewhere in the P's.

The following is an example of such a compound BON entry. It also illustrates the use of (A's and (B's. It is shown with the minor formatting provided by the tape printout programme: serial numbers are attached and the headword appears only on the first line of an individual entry; its repetitions on all subsequent lines are suppressed; a blank line is inserted between entries.

```
2685 G ETEY (A          +N, _(GEN)
     M                  *DEATH
     P                  NG'AAG NAN ETEY SI KANKANAN.  "IT IS BAD TO
     Q                  BE TALKING ABOUT DEATH."

2686 G ETEY (B          +V, _P.  P: NA- (NATEY).
     M                  TO *D̄IE; *DEAD
```

```
2687 $ P'ATAY              ETEY
     G                     +N
     M                     ANY *SACRIFICE PERFORMED ON BEHALF OF THE
     N                     VILLAGE

2688 $ PATEY (A            ETEY
     G                     +V, _A B I. A: IN-/-OM- (+REV), B: -EN
     G1                    (PATAYEN), I: I-.
     M                     TO *KILL; TO *EXTINGUISH OR PUT OUT, AS A
     N                     FIRE OR A LIGHT; TO TURN OFF, AS A RADIO OR
     O                     A MOTOR.  SEE: 'ADEP (A).

2689 $ PATEY (B            ETEY
     G                     +V, _A.  A: -OM- (+REV).
     M                     TO BE EXTREMELY *PAINFUL
     P                     POMALEY NAN 'OLOK. "MY HEAD IS KILLING ME."
```

The main entry in this example is ETEY, with its two parts of speech, noun and verb, distinguished with (A and (B after the headword. Grammatical information such as part-of-speech is recorded in the G band. Below these two entries is a subentry for the derived word P'ATAY, in whose $ band appears the root word, ETEY. Finally there is another subentry, for the derived word PATEY, which also has two parts of speech and hence appears as two entries. Each of these also contains a $ band referring to ETEY.

As subentries became more numerous, the artifice of adding the root in the sort handle often interfered with the (A's and (B's, causing entries to be sorted into the wrong order. The individual 'entries' in the above example were in fact once sorted wrongly, though they are shown here in the correct order. With some careful adjustment of the sort-handle generation programme this problem was corrected. The proper place for words with subscripts (the $1's and $2's used to disambiguate homonyms) in all this had also to be attended to by adjusting the sort handle. Nevertheless the need for these added complexities was just one of the many indications beginning to appear, that we were definitely outgrowing the system. However by then the BON dictionary was nearing completion, and none of the other band format dictionaries, including the new ones (Yapese (YAP), Kosraean (Kusaeian) (KSR), Mokilese (MOK), and Ponapean (PNP)) that had been initiated during 1971 and 1972, seemed to need these features. Furthermore, we thought (wrongly as it developed) that there would be no further dictionaries to be processed, and since we were busy writing other special programmes for those currently active, we looked no further.

During 1971-72 preparations were being made to commit the large Trukese (TRK) dictionary file to computer storage for further processing and expansion. The structure of entries in this file was a little more complicated than that of the BON entries. Each word could have two or more 'tense-modes' and within each tense-mode there could be two or more 'senses'. Bands would have to be repeatable for each different tense-mode as well as for each sense. Certain bands at the beginning of the entry would need to be shared by all the tense-modes in the entries: they would not be assigned to any single tense-mode. To accommodate these structural requirements by further bending the old band format in the way we had done for BON would have been extremely awkward. The only solution was to design a new format for TRK.

In this new format, the letters that labelled the bands were prefixed with numerals to indicate the level and sequencing of the tense-modes. We made no attempt to keep the requirement of sortability of the file (requirement 2) and abandoned the headword field. Instead of being repeated on each line, the

headword was assigned its own band, labelled with a blank instead of a letter, and positioned at the beginning of the entry. Since this dictionary, instead of being keypunched, was being entered from a terminal directly into computer (disc) storage, where there was no question of cards being misfiled, the headword as a field on each record was neither needed nor, for ease of typing, desirable. The format was easy for the typist, and was not too hard to read (requirement 1) so that no special print-out programme was needed, at least for routine proofreading purposes. Addition of new types of information was not anticipated, so requirement 3 was not critical. Since the process of entering and proofreading this massive dictionary was to be a long-term project in itself, we were little concerned with the fact that none of our programmes for 'standard' band format files could be used on it. Programmes for this dictionary would not be needed until some indefinite time in the future.

Following is an example of a TRK entry, as typed in upper and lower case:

```
    arakak
  c 3
  e arakaka
1   n. (T3)
11a *ringworm and similar skin *diseases;
11f ef=ew a. |one r.;
11g arakakan |his r.;
11i arakaken cuuk |Trukese r.;
12a *candle_*bush (*%Cassia_alata|);
12d [the pounded leaves are used as medicine for ringworm
12d and prickly heat.]
2   vi.
21a have ringworm;
21i Raa a. |They have ringworm.;
```

The band names, single letters, are in the third column, preceded by digits for tense-modes and senses. The headword band and the tense-mode band itself (the first band of each tense-mode) have blanks instead of letter names. The body of each band begins in the fifth column, separated from the band names by a blank column for readability. The headword of this entry is arakak, appearing in the body of the headword band (first line in example). There are two tense-modes, n. and vi., and the first tense-mode has two senses, *'ringworm etc.'* and *'candle bush etc.'* Bands c and e (derivational class and root form) as well as the headword band belong to the entire entry and not to any of these subdivisions, so are not prefixed with any numerals.

The necessity of designing a special format for the TRK dictionary was another indication that the old band format was not sufficiently general. Other aspects of the system of programmes were also proving inadequate. Furthermore, the MOK dictionary was beginning to acquire subentries, and the KSR dictionary to acquire modes. When it appeared, in 1973, that more dictionaries were definately about to be computerised, we decided to design a more general format, and write new programmes for it, that would accommodate in one system as many of the requirements so far encountered as possible, while seeking simplification in generalisation. This decision was aided by the hardware and software improvements described in section 2.2.

The new format that evolved during 1973-74 after initial trials with the CHA, Woleaian (WOL), and PAL dictionaries can accommodate entries with more complex structures than heretofore encountered. An entry may now have multiple 'modes' such as the tense-modes in the TRK dictionary and the BON entries with

(A's and (B's. Modes may be nested to any depth: the TRK dictionary required only two levels, but the new format accommodates any number. There may be subentries, such as those in the BON dictionary, and each may have within it as much internal mode structure as a main entry. There may be subentries to any depth. Bands are no longer labelled by single letters or blocks of single letters. A band name may be made up of any short convenient string of letters (e.g. ETYM, DEF, or SYN). There may be more than one occurrence of a given band. If for instance there are three example sentences in an entry, each can be placed in a separate band labelled, say, SX.

The following two entries from the PAL dictionary, for the words oltobedechu'r and recho'rech, illustrate some details of how these structural features are keypunched or typed at a terminal. The period in the first column signals the beginning of an entry; two periods signal the beginning of a subentry. There are two subentries (for the derived words rechereche'l and merecho'rech) under the second main entry (recho'rech). Deeper subentries would be flagged by correspondingly more periods. The digits enumerate modes, and multiple digits (which do not occur in this example, but would look like those in the TRK example) indicate deeper nesting of the modes. The band name follows the periods or digits. Then, after a few spaces comes the body, or the information portion, of the band. A continuation line would have column 1 blank. The influence of the TRK format is evident but all vestiges of fixed fields have now disappeared, and headwords get a full-fledged band name.

```
.hw   oltobedechu'r
ma    oltobed a chur
lps   v.s.
ldf   *comical; *funny; *laughter-provoking
2ps   n.
2df   *joke
cf    chur$1, oltobed
sx    Ng kmal mle oltobedechur a nglikir.  |Their dance was very comical.
.hw   recho'rech
ps    n.
df    *stealing; *theft; *robbery; *selfishness
..hw   rechereche'l
ps    n.poss.
..hw   merecho'rech
lps   v.t.
ldf   steal (including wife, girlfriend, etc.); rob
2ps   v.s.
2df   selfish
3ps   n.
3df   thief
```

This input format was designed to be easy to keypunch or type in, but is the least readable of all the formats. Requirement 1 has been relaxed beyond recognition. For printing therefore, a programme was written which inserts indentations and blank lines to highlight the structure of the entry for the reader, producing what is known as an L-format (listing format) print-out, which is quite readable after some getting used to:

```
.hw  oltobedechu'r
           ma  oltobed a chur
             lps  v.s.
             ldf  *comical; *funny; *laughter-provoking
```

```
            2ps  n.
            2df  *joke
          cf  chur$1, oltobed
          sx  Ng kmal mle oltobedechur a ngliker.  |Their dance was
            very comical.

.hw  recho'rech
          ps  n.
          df  *stealing; *theft; *robbery; *selfishness

  ..hw  rechereche'l
            ps  n.poss.

  ..hw  merecho'rech
            lps  v.t.
            ldf  steal (including wife, girlfriend, etc.); rob

            2ps  v.s.
            2df  selfish

            3ps  n.
            3df  thief
```

Requirement 2 (direct sortability) has also been abandoned. Programmes are needed to generate and attach sort handles to the entries before any sorting can be done. However, since many small dictionaries are still kept on punched cards rather than on tape, and even most of those on tape depend on punched cards for input, the capability of properly replacing any card in a punched-card file is still desirable. Some keypunchers feel this need more than others. We have put this capability back into the punched-card format by allowing (but not requiring) a fixed-length field, of whatever length the keypuncher wishes, at the beginning of the card, into which may be punched any desired sequencing or identifying information — a serial number, headword, part of a headword plus serial number, etc. This field is ignored by the programmes. It is there only for the keypuncher's security. It is only a part of the card format and not, as the headword field used to be, a part of the information structure itself.

Requirement 3 (flexibility) has, on the other hand, been vastly expanded. Band names need no longer be limited to single letters, and as a consequence of lifting requirement 2 from band names they need no longer appear in alphabetical order in an entry, and may even be repeated. Modes and subentries, each of which may be nested to any depth, provide new dimensions of flexibility.

The WOL, PAL, and Itawis (ITW) (a Philippine language) dictionaries were the first to be entered directly in this format, in 1973-74, followed by the dictionary of Thompson, an Amerindian language of the Pacific North-west. The others still active when this format was implemented were converted during the following year from the old format: KSR, MRS, PNP, MOK, YAP. After it had been completely entered in the original format, the TRK dictionary too was converted, in 1976. Some dozen other dictionaries have been comfortably accommodated in this format.

A striking aspect of the dictionary projects was the reliance on the print-out itself — the paper on which it is printed — as the principal medium on which field data was gathered by the lexicographer. In most other computer applications a print-out is mainly consulted for the information it contains. But with the dictionary projects the print-out was used as much for writing on as for reading. Hence the physical characteristics of the print-out — quality and size of paper, type of binding, distribution of blank space on the paper,

etc. took on an unaccustomed prominence. The sparse print-out format is partly a response to this need. The programme can be instructed to leave even more blank space. It is not uncommon to hear requests for triple or quadruple spaced print-outs.

## 3.2. Editing

After storage and listing, the next most necessary facility for an active lexical file is an easy way to make changes to the file. If changes are extensive one might do better to repunch the entire file. This was in fact done for the MRS and CHA dictionaries while they were still on cards, after new fieldwork had resulted in a large amount of revision. For MRS, new band assignments were made and the system of transcription was changed.

More often however only relatively small revisions had to be made to a file. Revising a file stored on tape or cards is known as updating or editing. To update a punched card file, individual cards can be replaced with new cards. A tape file can be updated in a similar way, except that the replacement of individual records in the tape file must be done by a programme. The MRO tape was edited in this way, using serial numbers to identify the records to be replaced. (The programme actually sorted the new records in with the old according to serial number, and then deleted those old records that fell next to the new ones.)

The NUK file was at first edited in a similar way, by replacing entire records, redundant fields and all, on tape. After the non-redundant storage system was adopted, a new editing programme was written in PL/1 which allowed the user to address particular fields (e.g. headword, word class code, root, etc.) to be replaced in an entry. The definition however was often still spread over a number of fixed length fields, and to edit such a definition one had often to edit two or more of these fields. The resulting inconvenience and possibility of introducing errors partly accounts for the preponderance of short definitions in that dictionary.

When the first band format dictionaries (MRS and CHA) were put on tape an editing programme was written in Fortran that was oriented specifically to this format. Since the headword was repeated on each record of an entry, the programme was designed to permit the headword to be edited independently of the rest of the entry. An edit instruction gave either a replacement headword for an entry or a replacement for the body of a record (columns 21-80) in a particular entry. The location of the entry in which the replacement was to take place was specified by punching the serial number of the entry (entries were assigned serial numbers when placed on tape), and an individual record within an entry was located by punching, in addition, the one or two character subsequence (band name) field of the record. We could have chosen to use headwords rather than serial numbers for addressing entries but reasoned that a three or four digit number was easier to punch and less prone to error than a long word in a language perhaps unfamiliar to the keypuncher. Also, if two entries, by mistake, had the same headword, there would have been no way to distinguish them.

Both this and the MRO system of editing, as well as the systems used on the NUK dictionary, had the disadvantage that, in order to change anything in a record or field the entire record or field had to be repunched, and if the revised material spilled over the end of the record or field, either a continuation record had to be inserted to take the overflow or an existing continuation record had to be revised. As far as the amount of keypunching and the possi-

bility of introducing additional errors were concerned, these systems were little better than manual filing in a punched card file. Only the job of hunting through a file and replacing cards was turned over to the computer. Nevertheless, nothing better was available for several years until the new system of programmes was adopted, even though an experimental editing programme was written and tried out on the CHA dictionary.

The editing programme developed for the new system used a contextual scheme for specifying changes. The decision to use this scheme was based on three realisations:

1) One should be able to instruct the programme to replace only some specified portion of the material in an entry without having to repunch an entire record or field.

2) The unit within which this portion is specified should not be a record but an entire band — that is, such a portion would never straddle two bands but might indeed often cut across a record boundary.

3) The user should not have to worry about record boundaries at all when making changes to material in a band. Records are arbitrary units of storage, and need not correspond to the units of information that the user is interested in.

These considerations led to the conclusion that an entire band, however long or short, should, for editing purposes, be considered as a single unit, which we called a 'paragraph', no matter how it was stored. The band label should be an integral part of the paragraph. Any material in the paragraph should be editable merely by referring to the serial number, or other identifier, of the paragraph and by further specifying enough of the context within the paragraph. In the notation developed for the new edit programme, an instruction, say, to change SHARK to SHARKS in paragraph 2457 (serial numbers were now given to paragraphs instead of to entries or to records) would be punched as follows:

```
@2457.  /SHARK//SHARKS/
```

How the paragraph is broken up into lines on the print-out or on the original punched cards is of no concern to the person punching the edit instructions. In fact, on tape each paragraph is stored as a single variable-length record, a type of storage that SPITBOL programmes can handle even more simply than fixed-length records. Only on printing out are the paragraphs broken up as necessary to fit on a page. The editing programme can, in fact, be used on any file of variable-length records, not only dictionary files. The availability of string handling facilities in SPITBOL was especially convenient for writing an editing system to meet these new requirements.

This editing system was a considerable improvement over the old one. It nevertheless required the error-prone process of punching cards. After each edit run the new print-out still had to be carefully proofread against the old one on which the corrections had been pencilled. We had, however, been exploring the possibility of a potentially much better system, namely editing on-line at a terminal. Such a system has the advantage that corrections are made in the file as soon as they are typed, and the results can be immediately seen. Any errors made in the process become immediately evident and can be corrected. Another advantage for us is that, at the proper kind of terminal, upper-and-lower case material can be edited just as easily as all-upper-case material. When TSO gradually became more reliable we experimented both with entering and editing dictionaries on-line, even before our current system of programmes was designed. The final editing passes of the CHA dictionary, by then converted to upper-and-

lower case, were made on-line in a minimally altered, but still fixed-field, band format. Editing in this way turned out to be extremely awkward not only because of the fixed-length records but also because of the limited size of files that could be worked on with the available computer facilities. This limitation meant that we had to split the dictionary up into 20 or 30 files, and, because of space limitations on disc and the cost of storage, bring each of them individually in from tape as we were ready to work on it and copy it out to another tape afterwards, a process that required keeping careful records of the different files as well as frequent trips to the computing centre.

Two other MC dictionaries (TRK and PAL) were entered and edited at the terminal, principally because they were already nearly in their final forms and editing was to be minimal. The juggling of the large number of files into which they had to be divided for on-line access continued to be a major headache.

In sum, the promise of easy on-line editing of dictionaries has not yet materialised, due largely to the cost, limitations, and unreliability of TSO. In the meantime, punching edit cards remains our principal means of entering corrections and additions to dictionary files.

## 3.3. Finderlists

While information storage and editing facilities must be part of any active filing system, the generation of finderlists from dictionaries was the first specifically lexicography-related need that the linguists expressed, and one of the principal motivations for storing dictionaries in computers.

There were three projects, in fact, which needed the computer only for generating finderlists from already completed or published dictionaries. Two of these (the Eng.-Puluwat (PUL) and Eng.-Maori (MAO) finderlists) were begun when there already existed a rudimentary system of programmes for generating finderlists from band format dictionaries, but rather than use these programmes it was simpler and more economical to punch each of the dictionaries in a special-purpose format that could be directly sorted. The format for the PUL dictionary, for instance, was simplicity itself:

| | |
|---|---|
| PWE'L | BURNED |
| PWELE | SEA BASS |
| PWELE | FISH |
| PWE'N- | REASON |
| PWE'N- | BECAUSE |
| PWER | VINE |
| PWER | PLANT |
| PWER | HOT |
| PWER | GLOWING |

Note the repetition of the PUL word on several cards when it has multiple Eng. glosses. After sorting on the Eng., the sorted print-out was manually edited and augmented with additional Eng. entries, and then typed for publication. The Eng.-MAO finderlist, with a more complicated format, was edited and expanded into an Eng.-MAO dictionary in Auckland after having been keypunched here. The third and much later finderlist-only project, for Eng.-Rennellese (REN), was able to take advantage of the much improved and more flexible finderlist programmes described below.

To punch a dictionary file which was intended to serve other purposes than the one-time generation of a finderlist however, the rigid format with one Eng. word per card would have been unsuitable. The lexicographers generally wished to have these words automatically extracted by the computer from the normally phrased definitions in the dictionaries. Conventions for marking these words, and programmes to extract them, were therefore developed. Thus identification and extraction of these 'keywords', as they were known, became the first of a series of finderlist-generation tasks the computer was to be asked to perform.

In the MRO dictionary file for instance, words to be extracted were marked with an asterisk, as in the first sample entry in 3.1., repeated here:

KAKEP *METAL STRIP USED FOR BINDING*GO NEAR*SURROUND*ENCLOSE*CLOSE

From this entry finderlist entries would be generated which, after alphabetisation by Eng., would give

| | |
|---|---|
| CLOSE | KAKEP |
| : | |
| : | |
| ENCLOSURE | KAKEP |
| : | |
| : | |
| GO NEAR | KAKEP |
| : | |
| : | |
| METAL STRIP USED FOR BINDING | KAKEP |
| : | |
| : | |
| SURROUND | KAKEP |
| : | |

The device of marking keywords with an asterisk for extraction into the finderlist is still used in the present programmes, but many refinements have been added since its introduction in the MRO dictionary. One restriction of the MRO system was that, since the asterisk served also to mark the beginning of each definition, only the first word of a definition could be made a keyword. The word BINDING in the above example could not be starred. If a word within a definition was to be marked, the definition had to be turned inside out so that that word would appear at the beginning. What should have gone before the word was placed after the end of the definition, set off by two hyphens, as in these examples:

AMBAKOL *CROP--BUMPER*CROP--PRIMARY AND UNTOUCHED
AMBAQ *SLAVE*RANK--LOW SOCIAL

This awkwardness was partially reduced in the system used for the first band format dictionaries (CHA, MRS, BON) and for the NUK dictionary. These programmes recognised the period rather than the asterisk as the boundary of definitions. The asterisk could then be placed before a non-initial word of a definition. The programme would extract the text up to the next period on the right, and as far as ten characters to the left, of the asterisk. Identification of the boundaries of the definition, now no longer necessarily marked by asterisks, was a second task assigned to finderlist-generation programmes. This ten-character span to the left was long enough to encompass some of the phrases that most commonly preceded keywords, such as TO, TO BE, CAUSE TO, and A KIND OF. The asterisk itself was replaced, in the extracted definition, by a blank. The ten-character limit was merely a function of the finderlist print-out format,

which had a fixed-width field to the left of the keyword, rather than a structural constraint. It could have been made longer, but seemed to take care of most cases. Longer pre-keyword phrases were simply truncated. With these conventions, the above MRO definitions could have been phrased more naturally. The entries, had they been punched in band format, would have looked like this:

```
A AMBAKOL          BUMPER*CROP. PRIMARY AND UNTOUCHED*CROP.
A AMBAQ            *SLAVE. LOW SOCIAL*RANK
```

The corresponding finderlist entries generated by the programme would have looked like this:

```
    BUMPER CROP                 AMBAKOL
 UNTOUCHED CROP                 AMBAKOL
           :
           :
LOW SOCIAL RANK                 AMBAQ
           :
           :
         SLAVE                  AMBAQ
           :
```

Note that the keywords (CROP, RANK, SLAVE) are alphabetised and vertically aligned for ease of scanning, and that 'PRIMARY AND' has been truncated from the left of 'UNTOUCHED CROP' due to excessive length.

This print-out format was used for finderlists generated from the early band format dictionaries (MRS, CHA, BON) and NUK. The rigidity of the pre-keyword field soon became annoyingly restrictive, however, as authors more frequently needed to contort their definitions to protect them from decapitation. Allowing a wider pre-keyword field would only have alleviated but not solved the problem. There was also the consideration of readability of the output and the efficient use of space. With the capabilities of a computer available the finderlist should, it was felt, be able to look more like a dictionary than a traveller's list of ready equivalents or a keyword-in-context concordance. The finderlist began to be viewed more as a final product and less as a skeleton to be fleshed out by the linguist before publication. Hence making a finderlist 'presentable' was a third type of finderlist-related need that developed.

A format was finally worked out with the authors of the NUK and BON dictionaries which solved the problem and was to prove more flexible, susceptible to refinement, and acceptable to all the other dictionary compilers. Its primary innovation was to separate the keyword from the definition in which it occurred. The keywords would stand, in alphabetical order, at the left of the page. Indented under each keyword would appear all the definitions in which it had occurred. Each definition would be followed immediately (rather than at a fixed column) by the word or words in the source language whose entries contained the definition. This format overcame the necessity for a keyword to occur near the beginning of a definition. It also made it possible to have more than one keyword per definition. Following are two sample entries from the Eng.-BON finderlist in an early version of this format:

```
PAINFUL
     PAINFUL:  OT-OT; SAKIT.
     TO BE EXTREMELY PAINFUL:  PATEY
     TO HAVE A PAINFUL BACK:  EDEG.

PALATE
     PALATE:  TA=NGEN.
     THE SOFT PALATE:   IL-ILOK.
```

In a subsequent revision of the format we eliminated those lines on which the definition consisted only of the keyword, moving the source word(s) onto the same line as the keyword itself, as in the following example:

```
PAINFUL:  OT-OT; SAKIT.
     TO BE EXTREMELY PAINFUL:  PATEY.
     TO HAVE A PAINFUL BACK:  EDEG.
```

Since this format relaxed certain constraints on the phrasing of definitions, the dictionary authors now had to rephrase many of their definitions that had been contorted to satisfy the conditions of the old format.

Another such widespread revision was necessitated by a change in the convention for delimiting definitions in the dictionaries. The period had been used to terminate definitions, but this usage began to conflict with the ordinary convention of terminating abbreviations with periods. While we had once agreed with the authors that such conflicts would be rare, as the dictionaries increased in volume the occasions when a definition in the finderlist was prematurely truncated at the end of an abbreviation became distressingly frequent. It was therefore decided to reassign the function of terminating definitions to the semicolon, a symbol that had not been available on the 026 keypunch, on which the earliest dictionaries had been entered. The stretches of text between semicolons were now called 'phrases'. As a consequence of this change, several large dictionaries had to be carefully edited to replace certain periods with semicolons. This was typical of many details that had to be attended to when dictionaries grew larger and provided more opportunities for conventions to conflict with each other.

Encouraged by the string-handling facilities of PL/1 (our first finderlist programmes were written in Fortran, before the days of PL/1) and later of SNOBOL4, and even later by the efficiency of SPITBOL, we added in successive rewritings of the finderlist programmes many further refinements of the three types mentioned, in response to various needs voiced principally by the compilers of the MC dictionaries. For instance, while originally the keyword was considered to be everything from the asterisk up to the following blank, conventions were later made available for the user to specify easily exactly where a given keyword was to end if not at the next blank. Also, asterisks were now no longer replaced by blanks. They were closed up. This meant that an asterisk could be placed in the middle of a word, as in UN*COMMON, making a part of the word (COMMON) the keyword. Another useful feature was this: if some desired keyword could not be worked into the definition at all it could be placed in some other band designated for such words. This band could easily be suppressed before publication. Along with the regular definition band, this band could be searched by the finderlist programme for keywords and phrases. It was also permitted to disambiguate homographous keywords with subscripts (e.g. *LIE$1 and *LIE$2 as in *'tell a lie'* and *'lie down'*) to force different finderlist entries to be generated for them and for the phrases in which they were embedded. This device was needed so that instead of the single entry

```
LIE
     TO FOOL, LIE, CHEAT:  MIS
     TO LIE DOWN HERE AND THERE, LIE RANDOMLY:  GAILETAB
     TO LIE, TELL A LIE:  GATTIG, RUT
     TO SETTLE, LIE STILL:  WOLOPAG
```

one would be able to force two entries to be generated:

LIE
TO FOOL, LIE, CHEAT: MIS
TO LIE, TELL A LIE: GATTIG, RUT

LIE
TO LIE DOWN HERE AND THERE, LIE RANDOMLY: GAILETAB
TO SETTLE, LIE STILL: WOLOPAG

A final example of finderlist features: The current programme generates the finderlist in the standard dictionary format (with 'keyword' and 'phrase' bands) so that it too, like the dictionary proper, can be edited, independently of the dictionary.

The thrust of these and numerous other devices accumulated over several revisions and expansions of the programme has been to allow the lexicographer to control the content and form of the machine-generated finderlist with as little effort as possible beyond the normal writing of definitions. We have apparently not stopped introducing new devices. The finderlist programme is now the most expensive of the various programmes to run on the computer.

To summarise, the finderlist-related needs have followed the following trend: the machine-generated finderlist was originally thought of as a starting point — a first draft that would be edited and expanded into a work in its own right (as was the case for the three finderlist-only projects mentioned above). But it came increasingly to be regarded as a final product, for which much effort was spent in programming for new conventions and in refining the phrasing and punctuation in the individual definitions. This development was at least partly motivated by the fact that, under the pressure of time, both parts had to be worked on simultaneously. It was simply not feasible for most of the projects to wait until the dictionary was completed before starting work on editing the finderlist. In any case the finderlist has always been considered only to be a kind of wordlist or index and not a full-fledged dictionary in its own right.

## 3.4. Systematic generation of possible words

Dictionary compilation projects for languages that already possess a large amount of written material often use computers to search texts for words to be added to the dictionary and to print out their contexts for studies of their meanings and use. None of the dictionaries processed here has required concordances of texts to aid compilation. (However, a number of concordances of the illustrative sentences have been made after a given dictionary was completed, these sentences being the most extensive corpus of text assembled to date in the particular language.) However many of the projects have used another computer technique, first implemented by Carroll (1966) for collecting NUK vocabulary. In this method, the computer is programmed to print out all phonotactically possible words in the language up to some specified length. This print-out is then sifted carefully by literate speakers of the language for actually occurring words — usually only a very small percentage. For languages with large or even medium-sized phoneme inventories the print-out of all possible disyllables often runs into hundreds of thousands if not millions of forms. A print-out of monosyllables is, of course, much more manageable, and frequently quite productive of new words in the early stages of word gathering. Monosyllable and disyllable print-outs were frequently made for MC and other dictionary projects and for projects which were otherwise not computerised. Programmes for producing these print-outs have been written to order for each language in PL/1.

This allows us to take maximum advantage of the simple devices in PL/1 for generating and printing such regular patterns.

## 3.5. Unconventional characters

Linguists using computers often feel confined by the limited set of characters that can be keypunched and that can be printed by the computer. Fortunately most of the languages under discussion can be written with the Latin alphabet plus a few of the additional symbols available on the keypunch, especially since most of the languages have practical orthographies designed to be usable on the typewriter. Requirements for non-printable characters have been handled in different ways as they arose. Digraphs and trigraphs have been used to represent non-Latin letters. Since on the keypunch it is not possible to overstrike a symbol, a diacritic such as a stress mark has been represented by some character such as an apostrophe punched either before or after the base letter. Similarly, to mark a letter as a capital in a field or band that is considered normally to be in lower case we have punched a preceding # sign, e.g. #AMERICA. On an upper/lower case terminal this convention is of course not needed. When a string of phonetic symbols was to be represented, as in giving the pronunciation of a headword, we delimited the entire phonetic string in some way, typically by placing it in a band of its own, and established special correspondence rules, valid only within these limits, between keypunch and phonetic symbols. For typesetting purposes, font changes have been signalled by keypunch symbols adopted for this purpose: % for shift to italic, ! for roman, etc. The reader will no doubt have noticed instances of the above conventions in the examples that appear in the preceding sections.

Most authors were content to live with these conventions while working with their print-outs. As a dictionary approached the publication stage however, these conventions felt increasingly obtrusive since the print-out now represented more and more directly the final published book. Proofreading for accuracy of such details as capitalisation and font shifts became more intensive during the final editing phases and was made difficult by the representation of these typographical functions by special symbols.

Some of this difficulty was alleviated by writing special programmes to print out the keypunched dictionaries in upper and lower case. (Those entered at the terminal were of course already in upper and lower case.) But for cost and other reasons this was never done on a regular basis. A further refinement was to compose the entries so as to approximate the shape of dictionary entries in a printed book, and print the file in upper and lower case with overstruck diacritics and with multiple impressions for simulating boldface letters. This was done for CHA and BON by special programmes, but was not repeated because of cost.

For some languages the poverty of the keypunch font may be a serious deterrent to using the computer. This was the case for certain North North-west Amerindian languages until the arrival of a time-sharing system allowed us to use a Selectric terminal, equipped with a special element, as the input and output device. This is still not a completely satisfactory solution to the character set problem, because the printing speed of the terminal is extremely slow. It is not even a general solution since it depends on the existence of a specially designed element. Any general solution in the current status of the technology will probably carry a high initial price.

## 3.6. Non-English alphabetical order

The standard alphabetical order for most of the languages under discussion is different from that of Eng. A simple example is MRO, in which Q falls between K and L. The computer, however, sorts in Eng. order. The expedient used to sort MRO words into proper order was to punch K as J, which was not otherwise used, and to punch Q as K. After sorting and before publication the K's were replaced by Q's and the J's by K's. (The MRO examples in previous sections were shown in the correct spelling, not as punched.)

For the other dictionaries the linguists were mostly content to follow Eng. alphabetical order in their working print-outs. In fact, not all the languages had adopted standard orthographies or alphabetical orders when work on the dictionaries began. The early dictionary files were therefore simply sorted directly on their headword fields, in Eng. order. After official ordering rules were adopted, or in some cases changed, for a given language, the lexicographer would want the dictionary sorted according to the new rules for taking out to the field or, later, for publication. For such sorting, a general strategy was adopted that involved writing a programme to generate from the headword an additional field called a sort handle which, when sorted on, would yield the words in the desired order. In NUK for example, P follows B, T follows D, K follows G, and NG follows N. To generate the sort handle the programme would be instructed to make a copy of the headword, then replace any P's in it by BZ, any T's by DZ, any K's by GZ, and any NG's by NZ. Thus, to the entry for the word PEDIANGA would be added a field containing BZEDIANZA. After all the entries were sorted by their handles, the handles would be deleted before the entries were printed out. The actual headword did not need to be reconstructed as in the MRO case, because it was never changed in the first place.

The above is a simple example of constructing sort handles. Almost all the dictionaries required non-trivial sort handles, some much more complex than this, consisting actually of several handles in a row. Any alphabetisation scheme that is dependent only on the spelling of the word itself (and not on external information such as part of speech) can be specified by describing an appropriate handle-making transformation. The present sorting programme contains features to help the user experiment with different handle-making transformations. Again, SPITBOL is particularly well adapted for writing programmes to construct sort handles. The ability of the computer to alphabetise thousands of words accurately, especially in an unusual order, has often been cited by the lexicographers as a significant benefit of computer processing.

## 3.7. Checking and systematic alterations

When a compiler's attention eventually turned from expanding the dictionary to preparing it for publication (but frequently before this stage), a distinct new phase of activity typically began which made special demands on the computer. The most characteristic activity in this phase was checking for consistency at all levels, (e.g. in punctuation, abbreviations, and typography), and for parallelism of phrasing in different definitions, etc. Inconsistencies might have crept in when work was resumed after months or years of inactivity on the dictionary, or when two or more people had independently written material for the dictionary, or merely when the linguist was too busy writing an entry to attend to some unrecorded convention invented for a forgotten entry. It is probably desirable that the author not be burdened with such details while working on

recording the essential information, but whereas in a large and well organised lexicographical operation such as those behind the standard reference dictionaries of Eng. a separate staff is assigned to enforce the batteries of editorial conventions, and even a smaller book published in the usual way is subjected to blue pencil of the copy editor, our dictionary authors found themselves solely responsible for such details and furthermore for the consistent insertion of typographical signals, another chore normally performed by an editor. The accuracy of typographical signals was even more important than that of spelling and punctuation since an error in, say, a font shift can generally have much more conspicuous consequences in the printed book. This phase of a dictionary project has tended to be relegated to the end, and there it has often unexpectedly assumed the proportions of a major project in itself, stretching over many unanticipated months and years.

However, special programmes were helpful, not only in presenting the data in ways in which inconsistencies could be more readily spotted, but sometimes in rectifying an entire class of inconsistencies. In fact, as an author grew to appreciate the power of the computer in ferreting out inconsistencies, he would often be able to formulate additional and more refined tests for the computer to apply, for discovering ever more obscure inconsistencies. By making the checking easier, the computer actually generated a greater need for checking on the part of the lexicographers. This cycle often continued until some arbitrary halt had to be called.

Following are some typical kinds of inconsistencies: inconsistent spellings or misspellings of words (judgment/judgement, or triggerfish/trigger-fish) in different entries, unparallel phrases where no real difference in meaning is intended (a kind of/kind of/type of), inconsistent abbreviation (esp./especially, or sp./species), failure to mark on keypunched material unpredictable capitalisation, errors of font and capitalisation in Latin binominals, etc.

While such inconsistencies could in principle be discovered by relentless proofreading, the process was aided by special programmes of varying sophistication. The simplest programmes printed out all occurrences of a given band, say all definition bands. While scanning such a print-out one's attention could be limited to the conventions applicable to a single band, so that one was much more likely to notice anomalies than when presented with entire entries. More sophisticated programmes were more selective, and presented the data in special formats, perhaps in upper and lower case, making certain errors easier to spot. If one cared to formulate rules, they could be programmed and used to discover non-conforming material. The finderlist itself, of course, often served as a consistency-checking device of a special kind. Many inconsistent spellings and misspellings of keywords and unparallelisms of phrases were detected by scanning a finderlist.

Another way in which a print-out of a subset of a dictionary has been helpful is in the checking of specialised technical information. For instance, if the entries have been appropriately tagged, such as by some special symbol in a certain band, all the botanical terms, with their definitions, could be printed out and sent to a specialist for checking. This was in fact done for several dictionaries. Retrieval for research rather than checking purposes belongs in the discussion here, though it was done relatively rarely. It has so far been limited to printing out lists of words satisfying given grammatical or phonological criteria, e.g. all intransitive verbs, all reduplicated words, all words with geminate consonants, etc. Phoneme concordances of headwords were made for several dictionaries and proved to be highly useful for phonological and morphological studies. Word concordances of illustrative sentences in several dictionaries were also made, for research as well as for checking purposes.

Having discovered inconsistencies, mistakes, or infelicities, it was necessary to make corrections in the file. This could be done in the usual way, by editing the file. If however there were many corrections of the same kind to be made, and the environment of the correction could be exactly specified, a programme could be written to do the job. For instance, all occurrences of ENG. in the etymology band could be changed to ENGL., or all instances of KIND OF in the definition band that lacked a preceding A could have the A added.

Such programmed alterations were not necessarily restricted to the correction of errors. If an author decided to restructure the material in every occurrence of some band according to a specifiable scheme, this could be done by programme. One of the more frequently occurring examples of this kind of restructuring was occasioned by spelling reform or standardisation. When new spelling conventions were adopted, all the source language words already in the dictionary would have to be respelled. The programme would be instructed to find not only all the headwords but also all the other occurrences of source-language words embedded in example sentence bands, cross-reference bands, etc. For MRS, the conversion of tens of thousands of phonemically transcribed forms to their orthographic shapes required testing and applying an ordered set of some 50 context-sensitive rules, a process that would not have been attempted by hand, or even by programming in a language less powerful than SPITBOL.

This subsection has ranged over a wide variety of needs and programmes that are characterised by these properties: they are a miscellaneous class — they do not obviously fall into any of the major categories of programmes described previously; each is a special-purpose programme, written for a particular need and used typically only once; they have to do either with retrieving some specified subset of the material in the dictionary and presenting it in some special way or with modifying some subset according to specified rules. These programmes have to be written for each occasion, and at one time they accounted for most of the programming effort. In the new system we have simplified the process somewhat by providing two or three fixed programme frameworks. The specific retrieval or restructuring instructions for each situation still have to be programmed, but they can be inserted into one of these frameworks. This method has considerably expedited certain types of one-shot operations.

We also now have a broad-spectrum retrieval programme which has proved extremely useful for a variety of purposes, including checking. It is simply a closely printed listing of all, or specified, bands grouped by band name and alphabetised on the data in the band. It has the useful property that it can be run without any deliberate forethought as to what material should be retrieved, yet the format of the output easily leads the eye to detect inconsistencies and errors. It can also serve as a crude retrieval device that suffices for many different purposes.

## 3.8. Publication

After a dictionary in computer form had been worked over and proofread many times it was natural to look for some means of automatically preparing photo-ready copy directly from computer tape in order to bypass another manual step, that of typesetting, with the need for yet another proofreading pass.

The MRO dictionary was, in fact, photographed directly from a specially formatted, all upper-case, computer print-out. For the BON dictionary a special print-out in galley format was produced, using the newly available upper/lower

case TN train with diacritics, and with multiple impressions to simulate boldface type. Italics and other fonts were, however, not available, and print impression was not perfectly clear, so the publisher decided to retype the entire print-out with Selectric elements for photo-offset. A third, smaller, dictionary of ITW was also retyped for publication from a computer print-out.

The NUK dictionary was the first to be typeset from computer tape at one of the newly formed firms specialising in computer-controlled photocomposition. Although working out the technical and administrative details leading up to the successful composition of this dictionary stretched over several years and took an inordinate number of man-hours and reams of correspondence, the difficulties stemmed not from the technology itself but from the poor communication between the six or seven levels of technical and administrative personnel involved in a novel distribution of responsibilities. The purpose of saving the author another proofreading pass was nevertheless achieved, though the other theoretical advantage, that of speed, over the conventional typesetting methods, was not in this case realised. The composition of the KAP dictionary was helped by the previous experience, but since the technical details had to be worked out anew with another firm, the operation still was far from efficient.

The CHA dictionary was composed by yet another company, but fortunately we have been able to stay with this firm for all the other MC dictionaries subsequently published. These dictionaries have been able to take advantage of a general system of programmes and procedures that developed out of the experience gained with the first three. Specifically, after the CHA dictionary was composed we worked out a set of basic conventions for all the photocomposition tapes subsequently to be sent to this firm, as well as an administrative routine for all the principals involved with each dictionary.

The conventions work as follows. The general appearance of an entry, and of a subentry, etc. in terms of indentations, vertical spacing, width of columns, size and style of type, is fixed. All the bands in each entry are strung into a single paragraph and printed with a hanging indent (see typeset example below). However, within the entry the author has complete control over typographic style on a band-by-band basis. Let us use the following two entries from the WOL dictionary as an example:

```
.HW   GESHAP
          BA   GASHAPA
            lPS   N.
            lDF   *BILGE AREA OF A CANOE.
            lPX   SEUW G. |ONE SECTION WHICH IS CUT OFF THE BILGE
              AREA OF A CANOE.

            2PS   VN.
            2DF   TO *CUT THE BILGE AREA OF.
            2SX   RE G. WA WE. |THEY ARE CUTTING THE BILGE AREA
              OF THE CANOE.
          SEM   *CANOE-PARTS:  BILGE AREA

.HW   -GIT
          BA  -GITI
          GR  VERBAL *SUFFIX
          DF  *SMALL, NOT *MUCH (IN QUANTITY AND SIZE).
          PX  MWONGOGIT |TO EAT A SMALL AMOUNT.
          ANT  -LAP
          SEE  GITIGIIT, FARIGIIT.
```

Instructions for typesetting are given in terms such as these: everything is to be converted to lower case unless otherwise specified; headword (HW) band in boldface; base-form (BA) band in italics, enclosed in parentheses, followed by a period; part-of-speech (PS) band in small caps, preceded by the mode number if there is one, in boldface, which is in turn preceded by an em-space; grammar (GR) band in roman, enclosed in square brackets (remove any asterisks); definition (DF) band in roman (remove any asterisks); phrasal example (PX) band begins in italics (the shift back to roman in mid-band is already signalled in the data by the | sign); sentential example (SX) band treated in the same way as PX band, except capitalise first letter of band and first roman letter; delete semantics (SEM) band (which was used only for finderlist generation); antonym (ANT) band in boldface, preceded by 'ANT.' in small caps; cross-reference (SEE) band also in boldface, preceded by 'CF.' in small caps.

These instructions apply to the entire dictionary. Each lexicographer is free to set such conventions for his own dictionary. It is unfortunately not possible, for reasons of time and cost, to make trial samples for evaluating the effect of different conventions. The decisions for each book have to be based solely on studying those dictionaries which have already been published. Once these decisions are made, however, the instructions can be encoded directly and simply into instructions for the programme that generates the photocomposition tape from the dictionary tape. The portion of the photocomposition tape generated from these two entries according to the above instructions would look something like this:

> ¤2¤Bgeshap¤R (¤Igashapa¤R).¤n ¤B1. ¤Cn. ¤Rbilge area of a canoe. ¤Ise uw g., ¤Rone section which is cut off the bilge area of a canoe.¤n ¤B 2. ¤CVN. ¤Rto cut the bilge area of. ¤IRe g. wa we. ¤RThey are cuttin g the bilge area of the canoe.
> ¤2¤B-git¤R (¤I-giti¤R). [verbal suffix] small, not much (in quantity and size). ¤IMwongogit, ¤Rto eat a small amount. ¤CANT. ¤B-lap. ¤CCF. ¤Bgitigiit, farigiit.

Each typographic function is signalled by a 'box code' consisting of the 'box' (¤) symbol followed by another character. ¤B means 'shift to boldface'; ¤I, 'shift to italics'. ¤C and ¤R are for small caps and roman, respectively. ¤n represents an en-space, which together with a following word-space makes an em-space. ¤2 signals the start of an entry: everything up to the next ¤2 is to be set as one paragraph with a 1-em 'hanging indent'. There are more than a dozen other box codes not illustrated here, plus certain box codes that represent special, e.g. phonetic, symbols that we have had made up.

This tape is sent to the photocomposition firm where it controls a machine which composes entire pages by 'drawing' the letters on special paper with an extremely fast and high-resolution electron beam. The two entries in our sample would come back on these sheets looking like this:

> **geshap** *(gashapa)*. **1.** N. bilge area of a canoe. *seuw g.*, one section which is cut off the bilge area of a canoe. 2. VN. to cut the bilge area of. *Re g. wa we*. They are cutting the bilge area of the canoe.
> **-git** *(-giti)*. [verbal suffix] small, not much (in quantity and size). *mwongogit*, to eat a small amount. ANT. **-lap.** CF. gitigiit, farigiit.

Each page would have two vertically justified columns of this material, complete with page headings and page numbers. These camera-ready pages are then used by the printer to produce pages for the book.

## 3.9. Accessibility of the programmes

The programmes used to be run exclusively by the programmers themselves. For a long time everyone believed that it would be too time-consuming and unprofitable to teach the linguists to run programmes for themselves, for several main reasons: each linguist needed at first to run programmes only at rare intervals, so would not get the frequent practice so necessary for learning and maintaining complex rote skills; the programmes were designed merely to run correctly, and were not 'engineered' for uninitiated users (this would have required much more programming effort); the intricacies of the computer's control language (Job Control Language) and its error messages required much training, experience, and reading of obscurely worded manuals, to understand.

As we gradually learned the idiosyncracies of our programmes, e.g. the kinds of mistakes human users are liable to make and their typical consequences on the computer runs, and as we continued to make the programmes easier for ourselves to use, the prospect of other linguists learning to use them became more promising. This situation actually was forced upon us when several of the MC dictionaries began to need more frequent service from the computer. We then attempted, as an emergency measure, to train some of the linguists to run the more frequently needed programmes, such as the editing and finderlist programmes. This was facilitated by several circumstances: the linguists concerned were willing to learn; having gained more confidence in our programmes we had collapsed some of the multiple-job programmes into single jobs, which required much less physical effort and bookkeeping to run; we had started to store and maintain the programmes themselves on the newly-available TSO disc files, which obviated the inconvenience of duplicating, distributing, and updating large decks of programme cards for everyone that needed them; and the computer itself had become more reliable.

Nevertheless, user involvement even after the training was still halting, due to the relative infrequency with which each linguist used the programmes and to the complexities of setting up the job cards. Further, some of the users we had trained then moved away, conducting their dictionary work by mail. Still, it was significant that those lexicographers who depended on others to run jobs for them experienced a greater degree of frustration in using the computer than those who ran their own even though the latter had to cope with quantities of technical detail.

Some of the experience gained by teaching lexicographers to run programmes was useful in designing the new system of programmes. These were still written under pressure of time, and accuracy was still more of a concern than user-oriented conveniences. However these programmes have required much less training to use than the previous set, and have been used by several linguists on their own after some initial instruction from already initiated users. The most glaring fault in the present system (aside from the difficulty of the job control language, which will probably always be with us) is the lack of self-contained and detailed documentation, though this is being slowly corrected. Most of the knowledge of how to run the jobs is still disseminated by word of mouth and by example, that is, by an experienced user showing a novice how to set up the job deck. As more and more lexicographers are attempting to run the programmes for

themselves the need for more adequate documentation and tutorial aids becomes more pressing.

The programmes themselves are of course still not finished, and never will be. Like any fieldwork project, this one too has had to be terminated or drastically curtailed for external reasons, and the results made available to others, even though much 'data' remains to be examined and more continues to be accumulated.

## APPENDIX:

## LIST OF PUBLISHED DICTIONARIES MENTIONED IN THE TEXT
(Updated to 1981)

| titles | authors | computer processing | publisher |
|---|---|---|---|
| *A Maranao dictionary* | Howard P. McKaughan<br>Batua A. Macaraya | 1965-1967 | University of Hawaii Press, 1967 |
| *Marshallese-English dictionary* | Takaji Abo<br>Byron W. Bender<br>Alfred Capelle<br>Tony DeBrum | 1966-1976 | University Press of Hawaii, 1976 |
| *Nukuoro lexicon* | Vern Carroll<br>Tobias Soulik | 1966-1972 | University Press of Hawaii, 1973 |
| *Chamorro-English dictionary* | Donald M. Topping<br>Pedro M. Ogo<br>Bernadita C. Dungca | 1967-1974 | University Press of Hawaii, 1975 |
| *Puluwat dictionary* | Samuel H. Elbert | 1970 (finderlist only) | Pacific Linguistics C-24, 1972 |
| *Bontok-English dictionary with English-Bontok finderlist* | Lawrence A. Reid | 1970-1973 | Pacific Linguistics C-36, 1976 |
| *Kapingamarangi lexicon* | Michael D. Lieber<br>Kalio H. Dikepa | 1971-1973 | University Press of Hawaii, 1974 |
| *Kusaiean-English dictionary* | Kee-dong Lee | 1971-1975 | University Press of Hawaii, 1976 |
| *Yapese-English dictionary* | John T. Jensen | 1971-1976 | University Press of Hawaii, 1977 |
| *Mokilese-English dictionary* | Sheldon P. Harrison<br>Salich Albert | 1972-1976 | University Press of Hawaii, 1977 |

| | | | |
|---|---|---|---|
| *Ponapean-English dictionary* | Kenneth L. Rehg<br>Damian G. Sohl | 1972-1976 | University Press of Hawaii, 1979 |
| *Trukese-English dictionary* | Ward H. Goodenough<br>Hiroshi Sugita | 1972-1980 | American Philosophical Society, 1980 |
| *Itawis dictionary with English-Itawis finderlist* | James A. Tharp | 1973-1975 | HRAFLex Books, 1976 |
| *Woleaian-English dictionary* | Ho-min Sohn<br>Anthony F. Tawerilmang | 1974-1975 | University Press of Hawaii, 1976 |
| *Palauan-English dictionary* | Edwin G. McManus, S.J., revised and expanded by Lewis S. Josephs and Masa-aki Emesiochel | 1974-1976 | University Press of Hawaii, 1977 |
| *The complete English-Maori dictionary* | Bruce Biggs | 1967-1968 | Auckland University Press, 1981 |
| *Dictionary of the language of Rennell and Bellona, Part 2: English to Rennellese and Bellonese* | Samuel H. Elbert<br>Rolf Kuschel<br>Toomasa Taupongi | 1975-1976 | National Museum of Denmark, 1981 |

Hsu, R.W. and Peters, A.M. "Computers and Micronesian dictionaries: a chronicle of systems-design fieldwork among the lexicographers". In Bender, B.W. editor, *Studies in Micronesian linguistics.* **C-80:1-36**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.1*

# RELATIVES, EQUATIVES, AND INFORMATION STRUCTURING

Roderick A. Jacobs

## 1. INTRODUCTION

This paper[1] is concerned with three major questions:

1. Why in many languages do topicalised sentences resemble relative constructions in systematic ways?
2. Why are the markers for both constructions so often proforms, specifically why third person singular pronouns?
3. What dimensions significant for grammatical description are suggested by these phenomena?

The data we will use for our discussion of these questions are drawn from several of the languages forming the so-called Trukic (TK) Continuum of nuclear Micronesian (MC) (Bender 1971). But the issues that arise have wider application.

## 2. SOME RELATIVE CONSTRUCTIONS FROM THREE LANGUAGE FAMILIES

Look, for example, at the following three sentences, each of which contains a relative construction. The languages they represent belong to quite distinct families — Austronesian, Indo-European, and Uto-Aztecan.

1. At-mwal ke-la WE i-a-weri-ir re-pwe-le mwongo
   *boy* PL-*that that I*-ASP-*see-them they-going-to eat*
   *Those boys that I saw were going to eat.*

   Saipan Carolinian (CRL)

2. *Those boys THAT I saw were about to eat.*

3. Iv-im-ep kikit-am PE? ne-tewi-ve pem-kwa?-pi pe?-miyaxwen
   *this*-PL-Past *boy*-PL *that I-see*-REALIS *they-eat*-IRREALIS *they-were*
   *These boys that I saw were going to eat.*

   Cupeño (Uto-Aztecan)

In TK sentences like 1. above, the form we, *that*, might be called a relative pronoun and assumed to be a proform which substitutes for a noun phrase coreferential with the head noun phrase of the relative, i.e. at-mwal ke-la,

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 37-51. *Pacific Linguistics*, C-80, 1984.

*those boys*. But if that were so, we might expect the plural pronoun form ke-we since the head noun phrase is plural. Alternatively the form we might be a subordinating conjunction homophonous with the third person singular demonstrative and, perhaps, having only a historical connection with the demonstrative. This would explain the lack of plural marking. Forms with a plural demonstrative after the head noun phrase are not acceptable:

4. *At-mwal ke-la KE-WE i-a-weri-ir re- we-le mwongo
*boy* PL-*that* PL-*that* *I*-ASP-*see*-*them* *they*-*going*-*to* *eat*

although it is possible to omit the we altogether.

The Eng. sentence (2) has the form 'that' after the head noun phrase. Like TK we, the form looks like the third person singular demonstrative pronoun. It is often referred to as a relative pronoun and has been analysed as a proform for a noun phrase coreferential with the relative clause head. But there are several reasons to reject such a characterisation. First, although there is a plural form 'those' for 'that' in other usages, this plural form cannot appear in the relative pronoun position:

5. **Those boys THOSE I saw were about to eat.*

Secondly, relative 'that' differs from true relative pronouns — 'who(m)', 'which', 'whose' — in that it cannot occur after a preposition:

6. **the boy to THAT I spoke . . .*

The WH forms, when objects of a preposition, may be fronted along with the preposition. Prepositional phrase fronting is a fairly general process not restricted just to relative clauses and WH questions. Thirdly, another non-pluralisable 'that' incapable of following a preposition occurs before noun phrase complement clauses. Since this form cannot be coreferential with a head noun phrase, it would seem more reasonable to consider both forms as really one, a marker of embedding. Considerations like these have led Downing, Bresnan, Stahkle, and others to reject the relative pronoun analysis, treating 'that' as a subordinating conjunction and relativisation as being of two kinds. In one, a WH form replaces the coreferential noun phrase and is fronted, replacing the embedding conjunction 'that'. Alternatively the coreferential noun phrase is deleted and the conjunction 'that' remains in its clause-introducing position. Of course, if 'that' is a conjunction, its failure to take plural marking after a plural head is hardly surprising. Once again we have a case of homophony and the coincidence is heightened by the fact that the TK and the Eng. forms seem to fill just the same coincidental roles.

However it also happens that the relative marker pe? in Cupeño has the same shape as the Cupeño third person singular demonstrative pronoun. Furthermore although the pronoun has a third person plural form pemem, *those*, this plural never appears in the introducing position for relatives:

7. *Iv-im-ep kikit-am PEMEM[2] ne-tewi-ve . . .
*this*-PL-PAST *boy*-PL *those* *I*-*see*-REALIS
**These boys those I saw . . .*

Nor can pe? take a postposition:

8. *I?i nawishmal PE-YKA ne-yawaywi-ve . . .
*this girl* *that*-*to* *I*-*spoke*-REALIS
*This girl to whom I spoke . . .*

If we reject a relative pronoun analysis for this Amerindian language, we might assign to pe? the status of subordinating conjunction. So once again homophony would have struck — and in just the same places as in the other two languages!

## 3. DIACHRONIC EXPLANATIONS

Obviously there is something rather suspicious about so many similar coincidences in unrelated languages. Why should just these homophonies occur and in just these places? We cannot escape into a diachronic explanation, for we must then face the diachronic question as to why it should be a third person singular demonstrative that became a conjunction. Certainly we can find no easy answer by looking at the one language of our sample for which we have earlier texts, texts going back over a thousand years. Anglo-Saxon ðe and ðæt turn out to provide no answers not available in modern Eng. except perhaps that ðe relative constructions are a little more like TK in one respect — they allow two proforms between the head and the relative, the first agreeing in number and case with the head, the second being invariant ðe, presumably a third person form.

But even if there were a good diachronic explanation available, this would not help us understand the present workings of the language. Children learn to use the two 'that' forms, the two 'we' forms, or the two pe? forms with the appropriate inflection or lack of inflection according to the grammatical contexts. We have no formulation that encapsulates the intuitive perception of the native speaker in this regard. We shall seek some clues to a possible explanation by looking in more detail at what happens in TK.

## 4. RELATIVE MARKERS IN TRUKIC

The TK languages each have a rich and flexible demonstrative system marking locational, temporal, and discourse perspectives for the entities involved. With relatives, demonstrative-like forms are used as introducers when needed. In fact, relative clauses most often have no introducer at all. In some circumstances this absence of marking leads to ambiguity. Thus:

9. Ekkewe semiriit re - sowu kkechiuw re - má mwittir
*that*-PL *children they* HABITUAL *cry* *they die early*

(Lagoon Trukese (TRK))[3]

is most likely to mean *Those children who habitually cry die young.* But, since there is a rule like the Eng. relative clause extraposition, the sentence could also mean *Those children who die young cry habitually.* Sentence 9. survives without a demonstrative-like form, in part because clear intonational differences distinguish the two senses. An ordinary demonstrative inflected for number, and/or a demonstrative-like form, becomes necessary where the speaker especially needs to mark off a clause as an identifying sentence rather than as a main assertion. An additional alternative is available in some special proforms that most often occur in relative constructions lacking an ordinary noun phrase head as well as in certain other constructions. These we shall deal with later in this paper.

The full range of proforms following the head noun of a relative includes not only the we and la forms already mentioned but many others. However, for those languages where demonstrative determiners do not precede the head noun, i.e. for all TK languages in our sample except TRK, only we or la, the invariant forms, can be the second of two proforms following the head. If these are not relative pronoun demonstratives but conjunctions, then we cannot draw upon the special features of the demonstratives in any direct way. The ordinary demonstratives we and la, and their plurals are unlike the others in that both are used to mark discourse referents known or already obvious to both speaker and hearer. Typically the demonstrative we marks something previously mentioned in the discourse or something not in the perceivable vicinity. The demonstrative la indicates something perceivable by both participants but often closer to the hearer, or, more rarely, something to do with future time. This shared knowledge feature is one which would also characterise the relation between a head noun phrase and its attributive relative clause. The relative clause is semantically a kind of predication concerning the referent of the head noun phrase, but the relation between the head and the attribute is assumed by the speaker to be known to the hearer. So, if this invariant we or la is a conjunction synchronically, it is another coincidence that the semantic relation is just that relation which is posited for simple noun phrases with a true demonstrative proform. An alternative is to concede that these invariant forms are proforms and to claim that they have a conjoining function. But then we will need to explain what we mean by 'proform'. The lack of number agreement and the fact that these proforms never precede an ordinary demonstrative belonging to the head suggest that they are external to the head noun phrase and may not refer to it, just as *that* in 10. below, if it does refer, does not refer back:

10. *If there be any amongst you such that he doubts these things . . .*

## 5. THE IDENTITY OF la AND we FORMS

What is in question then is the identity of this introducing element for relative clauses. It seems to us that the arguments against considering Eng. *that* a relative pronoun substituting for the coreferential noun phrase apply with similar force to the TK forms (and to the Cupeño forms also). Indeed there is one further parallel. In some of the TK languages there is also a la form which acts as a clause complementiser — in Ulithian (ULI) it introduces propositions that are presupposed true. This la cannot be pluralised, nor can it be co-referential with the head noun phrase since there isn't one. It cannot stand for a missing noun phrase in the embedded clause since none is missing. Both the relative la and the complementiser la precede embedded clauses, and both kinds of embedded clauses have a non-assertive force. That is, the relative clause is not *asserted* to be true of the head but is *assumed* true, while in ULI and in other TK languages complement clauses which are indeed *asserted* cannot take the la complementiser. Thus in sentences meaning 'I think that he is leaving', the complementiser could not be la in ULI but must be the reportive complementiser bo.

If the la and we forms are not the kind of relative pronoun described, i.e. not like the WH relative pronouns in Eng., are they then conjunctions? Like the adverb category in many grammars, the conjunction category is apt to be a wastebasket for items we cannot pin down more precisely. In Eng., Cupeño (and its sister languages), and the TK languages, the forms generally considered to be conjunctions are invariant. This is certainly true of 'that', pe?, and la and

we. And they clearly conjoin in the sense that they are interposed between a main clause and a subordinate clause. However, unlike other subordinating conjunctions in the languages, 'that' pe?, and la or we can often be omitted with little or not apparent effect on the structure or meaning of the sentence. Moreover, unlike the other conjunctions in these languages, they all happen — by coincidence, it seems — to be homophonous with third person singular demonstrative pronouns.

With these considerations in mind, we favour a third position. The pronominal resemblance noted for these forms is no coincidence — they are indeed pronouns. And the invariant third person singular shapes are also no coincidence; they arise from the fact that they refer to third person singular entities — embedded sentences. In other words la and we in TK are *cataphors*, forward-looking anaphoric pronouns used here to announce sentences embedded as noun phrases, just as the omissible 'it' in 11. announces a factive noun phrase complement:

11. *You won't regret IT that you signed over your property to me.*

In an interesting sense these cataphoric pronouns function as articles, the la and the we preceding embedded sentences but following simple head nouns.

## 6. ASSERTION VERSUS ASSUMPTION

Interestingly this seems to be the role of we (in its TRK form ewe) in 12:

12. Enaan áát EWE e -wopwut nenggin
*that*(deictic) *boy that he-hate girls*

But 12. is different in a significant way. The sequence beginning with ewe is not an attributive forming a complex noun phrase with enáan áát as its head. Nor is it the asserted predicate, as e-wopwut nenggin is the asserted predicate in 13:

13. Enaan áát e -wopwut nenggin
*that*(deictic) *boy he-hate girls*
*That boy there hates girls.*

(TRK)

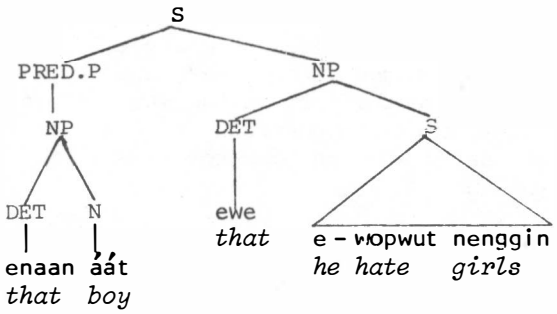
There is an important difference between 12. and 13. Sentence 12. is translated 'That boy there is the one that hates girls'. The speaker assumes that his hearer sees the individual referred to as enaan áát and he assumes that the hearer also knows already that some individual hates girls. The function of 12. is to identify the latter (the girl-hater) as the former, the boy over there. That is, enaan áát is the major predication, one which consists solely of a referential noun phrase. It is predicated of a referential noun phrase consisting of a sentence and introduced by the demonstrative determiner ewe. This sentence is, we claim, the TRK equivalent of the Eng. cleft sentence:

14. *It is that boy there that hates girls.*

As we shall argue later, the Eng. clause 'that hates girls', despite its appearance, is not a restrictive relative with 'that boy there' as its head.

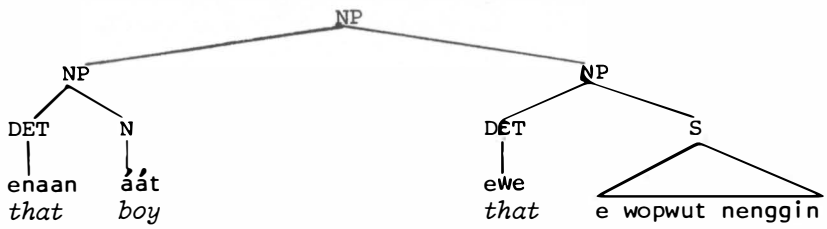
We might therefore show the relevant relations approximately as 15:

15.




It is of interest here that 12., with somewhat different intonation and, more often than not, with we rather than ewe, may serve as a noun phrase containing a relative clause. Such a noun phrase contains no assertion and so would be shown without a predicate phrase:

16.




The we form, like the enaan, does not obligatorily change its shape according to whether its noun phrase is asserted or purely referential and identifying in function. We are claiming here then that relative clauses with their heads form a single noun phrase consisting of two noun phrases, much like two equivalent noun phrases in apposition to each other. The second noun phrase, the less noun-like of the two, may have a determiner to identify it as a 'referring' sentence. But the determiner is optional.

So far, except for the so-called cleft sentence 12., we have examined only relative clauses. We have proposed an analysis in which two referential noun phrases are sisters forming a complex noun phrase in which the relation between the two major constituents is assumed (or 'presupposed') rather than asserted. We have argued that the optional introducing element for the embedded sentence is in fact a proform marked as singular because sentences when functioning as nouns cannot be plural. Hence the inability of 'it' and 'the fact' to pluralise in

17. $\left\{\begin{matrix}\textit{It is}\\ \textit{*They are}\end{matrix}\right\}$ *likely that he will resign and that a new government will be formed.*

18. *Is she aware of the* $\left\{\begin{matrix}\textit{fact}\\ \textit{*facts}\end{matrix}\right\}$ *that I founded the society and that I fund all its activities?*

## 7. TOPICALISED SENTENCES

We shall now turn to topicalised sentences. In 19. and 20. below iighi we *that fish* is the topic for both sentences, though its referent is in some sense the object of the verb angi:

19. Iighi we i-a-angi
*fish that I-*ASP*-eat*
*As for that fish, I ate it.*

20. Iighi we ILA i-a-angi
*fish that* TOPIC *I-*ASP*-eat*

While 19. has an intonational break after the topic noun phrase iighi we, 20. has the special form ila. Without the intonational break 19. could be used as a noun phrase containing a relative, *the fish that I ate.*

In sentences like 20. and those just below, the ila form follows the given information, separating it from the new information, which constitutes the major assertion of the sentences:

21. Ngaang YILAA yi sa tamolo
*I* TOPIC *I* ASP *chief*
*As for me, I've become chief.*

(ULI)

22. Ken INA e wairesin itto ikka
*Ken* TOPIC *he rare come here*
*As for Ken, he rarely comes here.*

(STW)

23. Sar mwal-la ye bel mwongo ILA semal sensei
*boy-*DEM *he will eat* TOPIC *one teacher*
*That boy who's going to eat, he's a teacher.*

(WOL)

The actual form ila resembles the form la that we have been discussing. It is similar in one other respect — it never occurs in plural form in this use. This may simply mean that it is merely a conjunctive particle. Since the relative determiner la has a we counterpart, it might be useful to check into the possibility of an iwe functioning like an ila. There is in fact an iwe. But according to Sugita (n.d.) in his study of TRK it is "used as a sentence adverb or conjunction with the meaning of 'then' or 'and then'". He adds: "It introduces an event that follows another event that is mentioned in the same sentence or in the sentence that precedes the one in which it is used." (5.32)

He gives examples like the following:

24. Ika kaa wocheey waach eey, IWE eni si-pwe mwootiw.
*if you-*ASP *eat-it boat-our this then probably we-will sink-down*
*If you (=mouse) eat our boat, then we will probably sink.*

(TRK)

We note here that both ila and iwe forms introduce a new assertion. In the case of the ila samples 20., 21., 22. and possibly 23., the sentences happen to look like relative clause constructions except that the internal clause is introduced by an ila (yilaa, ina) rather than a la (laa, na). In 20., 21. and 22. and their relative construction counterparts a noun plus a determiner is followed by a clause. In one case we would posit head noun phrases followed by relative clauses; in the other, topic noun phrases followed by comment clauses.

## 8. PREDICATIVE AND EQUATIVE SENTENCES

Now there are two major sentence types in TK languages. We will refer to them as *Predicative Sentences* and *Equative Sentences* though they have been given many other labels. Predicative sentences consist of a noun phrase subject, though it is sometimes omitted, and a predicate phrase beginning with a subject-referring proform known as a subject copy, and followed at least by a verb, perhaps preceded by negative and aspectual markers. Equative sentences have no main verb; both the subject and the predicate must be noun phrases. Thus 23. above is an equative sentence with the noun phrase sar mwal-la ye bel mwongo, *the boy who's going to eat*, as the subject and the noun phrase semal sensei, *a teacher*, as the predicate phrase. But embedded as a relative in the subject noun phrase is the predicative sentence ye bel mwongo, *who's going to eat*. *The* ye is the third person singular subject copy. In 23. ila separates the subject noun phrase not from a comment clause but from a comment simple noun phrase. We want to claim that comment clauses after ila are also noun phrases and that the -la part of ila is the cataphoric third person singular proform discussed earlier with regard to relatives.

Sentences without the optional ila, such as 25.,

25. Gaang semal sensei
*I one teacher*
*As for me, I'm a teacher.*
***I** am a teacher.*

(WOL)

are equative sentences. The predicate phrase has no subject copy and no verb. There is little point in treating 26. any differently,

26. Gaang ILA semal sensei
*I* TOPIC *one teacher*

(WOL)

since it has — minus a very slight emphasis — the same meaning, and has no subject copy and no verb. But sentences like 19. and 20., collapsed here as 27., present some difficulties for such a claim:

27. Iighi we (ILA) i-a -angi
*fish that* *I*-ASP-*eat*
*As for that fish, I ate it.*

(CRL)

The comment claise i-a-angi has both a subject copy and a verb. But if indeed ila includes a cataphoric proform, then the clause following it would be a sentential noun phrase, its internal structure irrelevant. For such a noun phrase the expected external subject copy would be the third person singular form e, marking i-a-angi as a predicate phrase in a predicative sentence. Such a combination is not permissible. However, the impermissability of this combination does not itself validate any noun phrase status for i-a-angi, even after ila, and this remains a weak point. Important here is the determination of the identity of the ila marker.

## 9. THE ROLE OF i-

Now in our discussion of relatives we claim that la, like we, is a cataphoric proform marking as a noun phrase the clause following it, unnounlike as that clause would otherwise seem. If indeed the -la of ila were this same cataphoric proform, what could we say of the i-? We note that without the i-, i.e. with just la, an equative relation between the two noun phrases is *assumed*, whereas with i-la the relation is *asserted*. The major difference lies in the asserted/presupposed relationship and, in TK, the i- marks this difference for the relation between noun phrases. We will mark ila as be-that in translations. We might say that the topic of any relative clause is the referent of its head noun phrase, i.e. topic is to comment as head noun phrase is to relative clause.

The i- then appears to be the introducer for an asserted predicate noun phrase, much as Gundel and Jacobs have claimed for 'be' and various existential presentative verbs in English, Thai, Russian, and other languages. The order Topic-Comment is not the only order, as our earlier example 12. might suggest. Ila (ina in TRK) is typically in initial position when the -la (-na) part itself is the only marking for the asserted noun phrase:

28. **I-NA** ewe iimw
*be-that the house*
*That one is the house.*

(TRK)

29. Ika **I-NA** wú-pwe feyinnó iye.
*if be-that I-will go-away there*
*If that's the case, I'll go there.*

(TRK)

If the asserted entity is animate, a marker is attached to a form of ila:

30. **I-LAALIR** mwal ke-la we re ghal kanta
*be-those man* PL-*that that they* ASP *sing*
*These are the men that were singing.*

(CRL)

31. **I-NAANIIY** ewe mwáán
*be-that-one the man*
*That is the man*

(TRK)

We will follow Sugita in calling ila forms and other i- forms 'predicative demonstratives'. Sugita's insightful discussion of TRK recognises the assertive characteristic of such forms and provides one especially illuminating contrast between a *this* form marked with i- and one without:

32. EEY efóch piin
*this one-long pencil*
*This is a pen.*

**I-YEEY** ewe piin
*be-this the pen*
*This is the pen.*

Such examples might be used to contrast ewe as a definite determiner with efóch, the long object classifier prefixed with the numeral e- *one*, as an indefinite determiner. At least as significant is the contrast between i-yeey and eey. Ewe piin in 33. is the nonasserted, i.e. given, information and the predicate introducer i- marks i-yeey as the asserted new information. But in 32. the lack of an i- indicates that the deictic eey is not asserted while the efóch-introduced noun phrase has to be the asserted predicate noun phrase.

The indications then are fairly clear that i-la, like the relative introducer la, is a cataphoric proform whose number depends on the constituent it introduces rather than the constituent preceding it. Where the introduced noun phrase is an embedded sentence, the ila cannot be marked for plurality nor, obviously, for singular animacy. Where the introduced information is referential and the reference is to animate entities, singular or plural proforms occur as appropriate. The ila announces sentential noun phrases as being noun phrases (the la role) and as constituting new information.

## 10. NON-ASSERTED SENTENTIAL NOUN PHRASES

Compare the following sentences, which contain much the same information though it is structured differently:

34. Iimw we I-LA i werii
*house that be-that I see-it*
*As for the house, I saw it.*

35. Iimw we MILLE i werii
*house that that-one I saw-it*
*It was the house that I saw.*

(CRL)

The content that is asserted in 34. is given in 35., and vice versa. Both 34. and 35. are equative sentences. We claim that the noun phrase iimw we is the subject of 34. while mille i werii is the noun phrase subject of 35. An interesting question then arises. Why would the asserted noun phrase, the so-called new information, have the we determiner? The answer is straight-forward enough. Sentence 35. is an identifying sentence. It identifies one entity already known to both speaker and hearer, 'the one (house) I saw' as being one and the same as another entity, 'that house' already familiar to speaker and hearer from other (unspecified) past experience. What is asserted is that the two known entities are in fact one.

The form mille (minne, menna, malaa, melwee, mele in other TK languages) marks an embedded sentence as a given non-asserted noun phrase. We might refer to these as me forms, since they may all contain a descendant of Proto-Micronesian (PMC) *me, a 'complete sentence particle' which survives as a relative clause marker in at least one other MC group, Ponapeic. This in turn may derive from a Proto-Austronesian *mea *thing*, which survives in Polynesian and other subgroups. The remainder of each of the mille forms consists of a demonstrative, suitably assimilated phonologically. The demonstratives are, not surprisingly, mainly the la and we forms. The me forms are non-assertive and seem to serve as both head and introducer, much as 'what' does in 'what he wants'.

These me forms, which we have seen in equative sentences such as 35., can also serve as a kind of relative marker in TRK. We have not found clear parallel examples in the other TK languages:

36. I-na ewe iimw MINNE osún we aa ássinó reen ewe tayifuun.
*be-that the house one roof-its that it-ASP blown by-it the typhoon*
*That's the house whose roof was blown off by the typhoon.*

37. Ewe áát MINNE itan Taro...
*the boy one name-his Taro*
*The boy whose name is Taro...*

(TRK)

In all the examples we are aware of, the me form occurs when there is a possessor/possessed relation that might be translated as *whose* for an Eng. relative clause. The me forms seem here to be head noun phrases in apposition to the immediately preceding noun phrase. Such a development is a likely one from sources like 38.,

38. Yato-we ye shiuweli melaw iye MELE-WE tamweoniu Molokai
*time-that he still alive he one-that chief-of Mologai*
*While he lived, it was he that ruled Molokai (he was the one that...)*

(WOL)

in which the construction beginning mele-we does not restrict the reference of the noun phrase iye *he* as a restrictive relative but is outside the noun phrase altogether. There is still the NP S structure usually associated with relative constructions and both iye and the mele-we construction refer to the *same* person. But the relation is asserted, while it would be assumed in a true relative construction.[4]

## 11. PARALLELS IN ENGLISH

We have claimed then that, structurally, TK relative constructions are noun phrases made up of two noun phrases, one containing the lexical head, the other consisting of a sentential description having the same referent as the head noun phrase. This complex noun phrase is the assumed (or 'presupposed') counterpart of an equative sentence in which the sentential description is asserted about the topic (corresponding to the head).

This is very much as if we considered Eng. constructions like 39. and 40:

39. *that doctor (who(m)) she married...*

40. *this house where I stayed...*

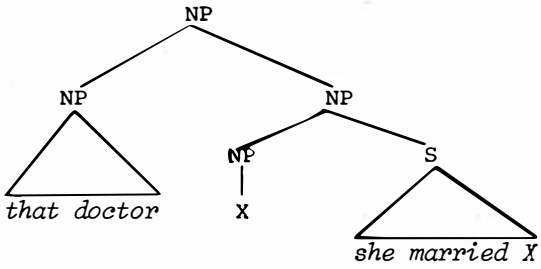
to be the non-assertive counterparts of the asserted equatives 41. and 42. below:

41. *That doctor is (the one) who she married.*

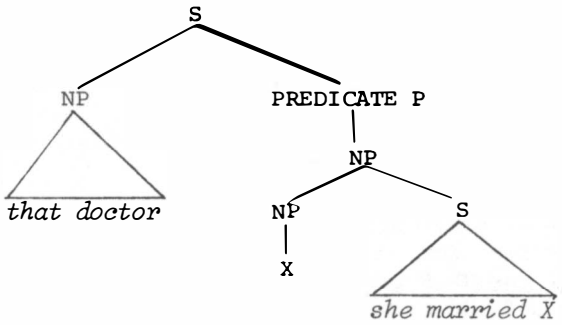
42. *This house is where I stayed.*

where the *is* verb corresponds to the TK i-. Examples 39. and 41. might be represented as 43. and 44. respectively:

43.




44.




Arguments for a double noun phrase analysis receive some support from cleft sentences like 45. and 46:

45. *It was Karolyn THAT I really liked.*

46. *The one THAT I really liked was Karolyn.*

In 45. the relative clause, which I claim to be a noun phrase, does not belong with its apparent head noun phrase, 'Karolyn', as the intonation also indicates. *That* relative clauses cannot have proper nouns like 'Karolyn' as heads. Moreover, the real assertion concerns the topic, some individual (x) that I really like. The assertion is that the individual was Karolyn. The apparent modifier of 'Karolyn' is, rather, an extraposed subject noun phrase, one required to have a lexical head, even if it is a false head, as is 'Karolyn' in 45., or a semantically almost empty head, as is 'one' in the unextraposed construction in 46. Downing, Stahlke and others have argued against relative pronoun status for 'that'. By 'relative pronoun' they mean a proform standing for a noun phrase (in a relative clause) coreferential with the relative clause head noun phrase. We consider 'that' a pronoun, but it is a cataphoric pronoun marking an embedded sentence as a noun phrase. In other words we do not feel compelled, after rejecting relative pronoun status for 'that', to accept simple conjunction status for this form. Our analysis means that the pronominal appearance of that, like those of its Uto-Aztecan and TK counterparts, is no coincidence but a consequence of the syntactic and informational context of the form.

## 12. SOME CONCLUSIONS

In our introduction we asked three questions. First we wanted to know why topicalised sentences in many languages resemble relative constructions. Our tentative answer is that semantically, both involve what we call equative relations between two noun phrases and that in both, the second noun phrase,

the sentential description, is *about* the first noun phrase. Many languages therefore 'economise' by using similar kinds of structures and forms, differentiating them according to assertive/non-assertive status. Secondly we asked why third person singular proforms were used in both constructions. We answer by saying that this is no coincidence. In both constructions the proforms are used to announce sentences as noun phrases, and sentences seem always to be third person singular noun phrases for those languages that have these distinctions. Finally we asked what dimensions significant for grammatical description are suggested by these phenomena. Perhaps the most significant dimension here beyond the more traditional syntactic and morphological criteria is that of informational structure, the structure that reflects the speaker's assumptions, genuine or hypocritical, about what his audience knows or believes relevant to what he is saying. Notions such as 'assertion', 'referentiality', 'topic', and 'comment' must be centrally involved in the formulation of the grammatical description of a language.

## NOTES

1. A shorter version of this paper was read at the Linguistic Society of America's Annual Meeting in Boston, December 1978.

2. However 7. is acceptable if the pemem is interpreted as itself a full noun phrase. There is, however, a slight intonational difference and, more conclusively, pemem can itself be followed by the non-pluralisable form pe?. This may also be true of the corresponding TK sentence 4. above.

3. The TRK examples here, as in most other parts of this paper, are drawn from an unpublished mimeo by Sugita. Although we have found many of his insights extremely valuable, he is not, of course, responsible for any extensions and possible misinterpretations by us.

4. TRK also has a meyi, described by Sugita as an aspectual stative marker typically but also used to mark what he calls "assertive mood". But meyi occurs initially in both 'stative' and 'dynamic' (non-stative) clauses, though it is obligatory only for the former type. It can also occur in sentence-initial position not only before the existential verb wor but before stative verbals, as in

> Meyi fókkun niyeyéech pwúnúwen ewe mwáán we.
> STATE *very woman-good spouse-of the man the*
> *The man's wife is very pretty.*

In this example meyi looks much like minne i.e. it may be serving as the head of a noun phrase in apposition to the noun phrase beginning pwúnúwen. It may be that the (y)i was the assertive i- discussed earlier. Hence the use of meyi for Sugita's so-called "assertive mood". If these speculations are valid, the main clause use of meyi developed from the embedded clause construction. Note also the similarity to the minne examples 34. and 35. of the following meyi sentence:

Enaan mwáán MEYI minefé newun ppwuk neyiy chóón sukun.
*that man new his book my person-of school*
*That guy whose book is new is my student.*

## BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

BRESNAN, Joan

1977 Variables in the theory of transformations. In Peter W. Cullicover, Thomas Wasow, and Adrian Akmajian, eds *Formal syntax,* 157-196. New York: Academic Press.

DOWNING, Bruce T.

n.d. English relativization. Mimeo. Department of Linguistics, University of Minnesota.

GUNDEL, Jeanette K.

1974 The role of topic and comment in linguistic theory. Doctoral dissertation, University of Texas at Austin. Reproduced by Indiana Linguistics Club.

GUNDEL, Jeanette K. and Roderick A. JACOBS

1980 Why "seem" and "be" aren't what they seem to be. *Centerpoint* 4/2: 115-123.

JACOBS, Roderick A.

1973 Syntactic compression and semantic change. In Claudia Corum, T. Cedric Smith-Start, and Ann Weiser, eds *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, April 13-15, 1973,* 232-241. Chicago Linguistic Society.

1975a Promotion and thematisation processes in English, or, How to get a head. In Robin E. Grossman, L. James San, and Timothy J. Vance, eds *Papers from the Parasession on functionalism, April 17, 1975,* 223-231. Chicago Linguistic Society.

1975b *Syntactic change: a Cupan (Uto-Aztecan) case study.* Berkeley: University of California Press.

1976a A passive continuum in Austronesian. In Sanford B. Steever, Carol A. Walker, and Salikoko S. Mufwene, eds *Papers from the Parasession on diachronic syntax, April 22, 1976,* 118-125. Chicago Linguistic Society.

1976b Syntactic reconstruction and the comparative method: a Uto-Aztecan case study. In Mohammad Ali Jazayery, Edgar C. Polomé, and Werner Winter, eds *Linguistic and literary studies in honor of Archibald A. Hill*, vol.1: *General and theoretical linguistics*, 165-174. Lisse: Peter De Ridder Press.

1977 Syntactic categories and predictions in Gilbertese. Paper presented to the Linguistic Society of America Summer Meeting, Honolulu, Hawaii.

1980 Towards a semantic grammar of English. MS.

KUROIWA, Kenneth

1975 BO and LA: factive and non-factive complementation in Ulithian. MS. University of Hawaii.

SOHN, Ho-min

1973 Relative clause formation in Micronesian languages. *Oceanic Linguistics* 12:353-392.

STAHLKE, Hervert F.W.

1976 Which that. *Language* 52/3:584-610.

SUGITA, Hiroshi

n.d. Lagoon Trukese grammar. MS. University of Hawaii.

Jacobs, R.A. "Relatives, equatives, and information structuring". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:37-51**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.37*

# THE ORIGINS OF 'COMPENSATORY LENGTHENING' RULES IN MICRONESIAN LANGUAGES

Kenneth Rehg

In studies of Micronesian (MC) languages, considerable attention has been paid to what most linguists working in this area have traditionally — and I believe mistakenly — called 'compensatory lengthening'.[1] The nature of this phenomenon is illustrated by the following examples from Ponapean (PNP).

| Base | Construct Form | Free Form | Gloss |
|---|---|---|---|
| *kili | kilin | ki:l | *skin* |
| *seti | setin | se:t | *sea* |
| *mwarɛ | mwarɛn | mwa:r | *title* |
| *ɔsɛ | ɔsɛn | ɔ:s | *thatch* |
| *roŋɛ | roŋɛn | ro:ŋ | *news* |
| *umwi | umwin | u:mw | *earth oven* |

The base forms listed in the first column represent, depending upon one's theoretical persuasion, either the synchronic underlying forms of these morphemes or historically earlier forms arrived at through internal reconstruction. The construct forms of these morphemes (meaning 'X of'; therefore, 'skin of', 'sea of', etc.) are derived by suffixing ni to the base and deleting the final vowel. The free forms of these morphemes differ from their corresponding base forms in two respects.

(1) The final vowel of the base is deleted.

(2) The first or remaining vowel of the base is lengthened.

Thus, all short base vowels are lost before word boundary in PNP. Lengthening, however, operates only on the first vowel of a disyllabic base of the shape (C)V(C)V. The construct forms of the morphemes do not exhibit lengthening, nor do the free forms of bases of shapes other than (C)V(C)V. Note, for example, the following forms.

| Base | Free Form | Gloss |
|---|---|---|
| *nsara | nsar | *snare* |
| *ɛmpi | ɛmp | *coconut crab* |
| *aramasa | aramas | *person* |

An excellent summary of the history of the study of such vowel lengthening in MC languages is provided by Bender (1973), who notes that the earliest account

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 53-59. *Pacific Linguistics*, C-80, 1984.

of this phenomenon was provided by Dyen. Dyen (1949:423) observes that the double or long vowels in the free forms of earlier disyllabic words in Trukese (TRK) can "be traced directly to compensatory doubling.... The last vowel was lost and the preceding only vowel of the word was doubled". Although Dyen (1949:423) does not make explicit which TRK word classes undergo compensatory doubling, all of the examples he cites are nouns. He does observe that denominal verbs as well as nouns in close fusion with prepositions do not undergo compensatory doubling "because any such word was part of a closely connected sequence, and the sequence was longer than a monosyllable". He gives these examples, using q for what is today spelled pw.

| | |
|---|---|
| qiin | *night* |
| qini-n | *night of* |
| ja a qin | *it is night* |
| nee qin | *at night* |

Bender (1971:440-441) notes that in TRK

> both consonants and vowels occur long (doubled and short, although some vowel length is inherent... and other vowel length is compensatory..., the result of a late stage rule requiring independent forms of some word classes to have the equivalent of two syllables.

Again, which word classes undergo compensatory lengthening (Dyen's compensatory doubling) is not discussed.

So far as I am aware, the earliest formulation of a rule that restricts this lengthening to a specific word class was provided by Irwin Howard at the University of Hawaii at Manoa. Using TRK data in problem sets for his phonology classes, Howard not only specified nouns as the word class this rule of lengthening operates on, but also demonstrated that, within a generative framework, a number of facts about TRK phonology could be more effectively accounted for if one ordered lengthening prior to final vowel deletion. What had previously been assumed to be compensatory lengthening was shown by Howard to indeed be, at least synchronically, anticipatory lengthening (Bender 1973:467-469).

My own work on PNP phonology in the early 1970's (Rehg 1973) mirrored Howard's analysis of TRK and ordered lengthening prior to final vowel deletion. The following three rules, informally stated and somewhat simplified for this presentation, exhibit the ordering of the PNP rule of lengthening in relation to final vowel deletion as well as to a rule of raising.

Rule 1: Lengthening

$$[\#(C)V(C)\#]_N \quad \Rightarrow \quad [\#(C)V{:}(C)V\#]_N$$

This rule lengthens the first vowel of an unaffixed disyllabic noun.

Rule 2: Raising

$$\begin{bmatrix} a \\ -long \end{bmatrix} \quad \rightarrow \quad \varepsilon \; / \_\_\_\_ (Ci)(C)i\#$$

This rule raises short /a/ to /ɛ/ when the final vowel in a word is a high front vowel, providing that no vowel other than /i/ intervenes.

Rule 3: Final vowel deletion

$$V \quad \rightarrow \quad \emptyset / VC(V) \_\_\_\_ \#$$

This rule deletes a final vowel before word boundary.

I will not attempt to justify the particular form of these rules. In fact, I am now reasonably confident that, as a synchronic account of PNP, they are in error. But, as the following sample derivations illustrate, it is clear that lengthening must precede raising (so as to block raising in certain forms) and that raising must precede final vowel deletion (since the rule of raising includes this final vowel in its environment).

| | | ina | ina+ni | laŋi | laŋi+ni | laŋi+na |
|---|---|---|---|---|---|---|
| | | *mother* | *mother of* | *sky* | *sky of* | *cloudy* |
| Rule 1: | Lengthening | i:na | -- | la:ŋi | -- | -- |
| Rule 2: | Raising | -- | inɛ+ni | -- | lɛŋi+ni | -- |
| Rule 3: | Final vowel deletion | i:n | inɛ+n | la:ŋ | lɛŋi+n | laŋi+n |

Lengthening, raising, and final vowel deletion rules interact in a parallel manner in TRK.

It has been assumed that lengthening in all MC languages which exhibit this phenomenon could be accounted for in essentially the same manner as it is accounted for in PNP. Yet, I think that all Micronesianists have in fact been troubled by this account of lengthening. Indeed, it is reminiscent of a cartoon I recently saw. In this cartoon, a mathematician is standing before a large blackboard that is filled with a long, complex proof. In the middle of this proof is the phrase "and then a miracle occurs". At least two aspects of this account of lengthening strike me as mysterious, if not miraculous. The first is, if lengthening has anything to do with final vowel deletion, then how could it literally be anticipatory? I do not believe that languages, any more than the people who speak them, have the ability to see into the future. The second is, why should this rule of lengthening apply only to nouns? It is obvious that lengthening requires further explanation and investigation.

With respect to the issue of rule ordering, it is clear that historically at least, lengthening did precede final vowel deletion. For example, in some western Trukic (TK) languages, base final vowels are retained as voiceless vowels; nevertheless, lengthening of the first vowel of disyllabic bases still occurs, as illustrated by the following examples from Woleaian:

| Base | Free Form | Gloss |
|---|---|---|
| laŋo | la:ŋo̤ | *fly* |
| yafi | ya:fi̥ | *fire* |
| ita | i:te̥ | *name* |
| mata | ma:te̥ | *eyes* |
| rigi | ri:gi̥ | *running* (noun) |

Ward Goodenough has also pointed out (personal communication) that, in Kiribati (Gilbertese) (KIR), the only base-final vowels that undergo deletion are high vowels after a nasal. Nevertheless, a similar pattern of lengthening exists in this language, as illustrated by the following examples.

| | |
|---|---|
| te mata | *eye* |
| matana | *his eye* |
| ma:ta | *eyes* |

The plural translation *eyes* for the form ma:ta indicates that when a noun is used without the article, or is in no other way modified, it is given a generic

reading. It should also be noted that KIR lengthening affects surface monosyllables as well as disyllables, as illustrated by the next examples.

| | | |
|---|---|---|
| te tona | to:na | *yaws* |
| te umwa | u:mwa | *house* |
| te ika | i:ka | *fish* |
| te pen | pe:n | *ripe coconut* |

But, as in TRK and PNP, KIR lengthening does not affect bases of greater length.

| | | |
|---|---|---|
| te pwapwai | pwapwai | *root of a taro-like plant* |
| te mwmwaane | mwmwaane | *man* |

Based upon evidence from western TK languages and KIR, it thus seems clear that, diachronically, lengthening did precede final vowel deletion. However, the issue of whether or not this lengthening has in fact anything to do with final vowel deletion is not resolved by these data. I will return to this issue later.

The order of lengthening with respect to final vowel deletion was one of the "mysteries" to which I referred earlier. The other was the fact that lengthening, as commonly described for TRK and PNP, is presumed to apply to nouns only. With regard to this concern, data from KIR are again particularly interesting, since it is *not* true in KIR that lengthening is so constrained. Note the following examples:

| | | | |
|---|---|---|---|
| E piri | *He ran.* | Pi:ri! | *Run!* |
| E nako | *He went.* | Na:ko! | *Go!* |
| E siku | *He stayed.* | Si:ku! | *Stay!* |
| E kipa | *He jumped.* | Ki:pa! | *Jump!* |

These forms illustrate that when a disyllabic verb occurs alone in a phrase, the first vowel of the verb undergoes lengthening. Verbs thus behave precisely as nouns do. Therefore, as with nouns, lengthening does not apply to verbs of three or more syllables, as demonstrated by these examples:

| | | | |
|---|---|---|---|
| E osinako. | *He went out.* | Osinako! | *Go out!* |
| E anene. | *He sang.* | Anene! | *Sing!* |

In KIR, therefore, lengthening is not constrained to apply only to nouns. A closer examination of PNP reveals that, in fact, lengthening is not so constrained in this language either. While verbs never undergo lengthening, other word classes do, as illustrated by PNP reflexes of the following Proto-Micronesian (PMC) forms.

INDEPENDENT PRONOUNS

| PMC | PNP | |
|---|---|---|
| *ia | i: | 3rd pers. sing. |
| *kit'a | ki:t' | 1st pers. excl. |

Since independent pronouns never occur with affixes or in combination with other words in a phrase, they do not exhibit alternations of length however.

NUMERALS

| | |
|---|---|
| e:w | *one* (independent form) |
| yew | *one* (enclitic form) |

POSSESSIVE FORMS

| | |
|---|---|
| a:y | *mine* |
| ɛy | *my* |
| a:mw | *yours* |
| omw | *your* |
| a: | *his* |
| ɛ | *his* |

QUESTION WORDS

| | |
|---|---|
| ta: | *what (What is this?)* |
| ta | *what (What kind is this?)* |
| i:s | *who* |
| isime | *who* |

What all of these word classes have in common, along with nouns, is that they may occur alone without further modification in a noun phrase. When alone, they have a long vowel; when not alone, they have a short vowel. Only words having the potential to stand alone in noun phrases in PNP exhibit this kind of lengthening.

The examples from KIR and PNP suggest that, diachronically, lengthening of the first vowel of a disyllabic morpheme was a *phrase* governed phenomenon, not a *lexically* governed one. Therefore, the origins of lengthening might be explained in terms of the phrase. Two related explanations come to mind.

The first is that in a language ancestral to PNP, KIR, and the TK languages (as well as possibly to other MC languages), a constraint existed to the effect that all phrases had to be minimally trimoric. Since all lexical items belonging to major word classes were apparently minimally disyllabic in this language (as in Proto-Oceanic), noun phrases were normally at least trimoric as a consequence of the presence of an article. Verb phrases, too, normally contained at least three mora because of the presence of a subject pronoun. When disyllabic nouns were used generically, with no further modification, or when disyllabic verbs occurred in commands, with the deletion of the subject pronoun, the constraint that phrases be minimally trimoric was satisfied by lengthening the first vowel of the disyllabic base.

The preceding account of the origins of lengthening is not entirely satisfactory, however, for the simple reason that it fails to explain why it is the *first* vowel of a disyllabic base that lengthens. Three mora could just as easily be arrived at by doubling the last vowel, by adding a prothetic vowel, by doubling a medial consonant, etc. The observation that phrases must have been minimally trimoric in some ancestral Micronesian language is better explained, I believe, as a consequence of the interaction of the following two quite natural rules (in which ‖ marks phrase boundary).

STRESS ASSIGNMENT

V → [+stress]/____(C)V‖

PHRASE INITIAL STRESSED VOWEL LENGTHENING

V → [+long] ‖ (C)____

The first rule assigns primary stress to the penultimate vowel of a phrase. Penultimate stress is the expected position of stress for Oceanic languages, and I have argued elsewhere (Rehg 1978) that one must assume such a stress pattern

for Proto-Ponapeic to account for certain vowel deletion phenomena that take place in Mokilese, a Ponapeic language. The second rule lengthens a stressed vowel when it is the first vowel in a phrase. Many phonetic arguments can be brought to bear in support of the naturalness of this rule, including the heightened air pressure that exists at the onset of a breath group, the common occurrence of phrase-initial stress in many languages, and the expected (but not inevitable) pattern of stressed vowels having greater duration than unstressed vowels. These two rules, then, might account for the genesis of the lengthening of disyllabic bases in MC languages. When the phrase was only two syllables long, the stressed vowel occurred in phrase-initial position and was thus lengthened. If the phrase contained more than two syllables, the stressed vowel was never in phrase-initial position, and was thus *not* lengthened.

The preceding discussion, if correct, renders vacuous the issue of whether the lengthening of disyllabic bases was anticipatory or compensatory with relation to final vowel deletion. Such lengthening quite probably had nothing to do with final vowel deletion. In contemporary KIR, for example, it is not the loss of a final vowel which triggers lengthening, but rather the position of a stressed vowel in relation to the beginning of the phrase. Thus, a surface monosyllabic form will contain a short vowel in a phrase like te pén *the ripe coconut*, but will have a long vowel in a generic phrase such as pé:n *ripe coconuts*. It is presumably the case, therefore, that in modern KIR stress is assigned prior to final vowel deletion, or the rule governing stress has been modified to the form $V \rightarrow [+\text{stress}] / \_\_\_\_ (C)(V) \|$ . The presence or absence of final vowels in surface forms plays no role in determining when lengthening applies.

While lengthening in KIR can possibly be accounted for by only slightly modifying the two rules presented earlier, it is quite clear that these rules will not suffice for contemporary PNP nor for the TK languages. How synchronic lengthening is to be accounted for in PNP is a problem I am currently investigating, but for which I have not yet arrived at any satisfactory solution. Thus, the resolution of this issue must be considered at another time.

The final point I wish to cinsider here concerns the relative antiquity of phrase-initial stressed vowel lengthening. As Bender (1973) noted, the lengthening of disyllabic bases occurs in all TK languages, in KIR, and, I believe, must certainly have been a feature of Proto-Ponapeic. Two modern Ponapeic languages — Mokilese and Pingelapese — no longer exhibit such alternations in length, but evidence exists that they once did. Geoffrey Nathan has reported (personal communication) that sporadic lengthening exists in Nauruan, and Kee-dong Lee and Judith Wang have observed that all monosyllabic bases have long vowels on the surface in Kosraean, at least when those bases are uttered in isolation. So far as I am aware, only Marshallese fails to offer evidence for the existence of an earlier rule of phrase-initial stressed vowel lengthening. Whether or not this observation will prove useful in determining subgroupings within the MC family of languages remains to be seen, but it is a line of inquiry worthy of further investigation.

## NOTE

1. An earlier version of this paper was presented to the Diachronic Linguistic Festival at the University of Hawaii in November of 1979.

## BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

1973 Parallelisms in the morphophonemics of several Micronesian languages. *Working Papers in Linguistics, University of Hawaii* 5/8:1-16. Reprinted in *Oceanic Linguistics* 12 (1973):455-477.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

REHG, Kenneth L.

1973 On the history of Ponapean phonology. *Working Papers in Linguistics, University of Hawaii* 5/8:17-55. See also Rehg 1983.

1978 Toward a theory of Ponapean prosody. Paper presented at the Linguistic Society of Hawaii Fourth Annual Oceanic Linguistics Festival.

Rehg, K. "The origins of 'compensatory lengthening' rules in Micronesian languages". In Bender, B.W. editor, *Studies in Micronesian linguistics*.
**C-80:53-59**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.53*

# HIGHER NUMERALS IN SEVERAL MICRONESIAN LANGUAGES

Sheldon P. Harrison
Frederick H. Jackson

## 0. INTRODUCTION

In a number of Micronesian (MC)[1] languages, there exist monomorphemic numbers for multiples of ten, reaching $10^9$ in the case of Ponapeic (PP) languages. (In one Polynesian (PN) language, Nukuoro (NUK), a similar series reaches $10^{10}$.) Reports of the existence of higher number systems of this sort are often treated with a great degree of scepticism. A typical reaction is that of Elbert and Pukui (1979:160) who, in considering the Hawaiian (HAW) numbers lau, mano, kini, and lehu,[2] state that "they are used poetically as nouns indicative of great numbers".

They note (1979:161)[3] that "elsewhere these quantities are somewhat fancifully translated as 400, 4,000, 40,000, and 400,000. It is doubtful that actual counts of this magnitude were ever made.... Of even greater magnitude is nalowale, usually translated *lost* but sometimes considered a number equal to ten lehu, which is four million. It is inconceivable that people counted that many".

While it is undoubtedly true that there were no objects in any society that one would want to count into the hundreds of thousands or millions, we do not find it inconceivable that an abstract mathematics with linguistic means of representing quantities of such magnitude or beyond could have developed in traditional Oceanic societies. In our experience, speakers of MC languages will present the higher number series in order [$10^1$ through $10^x$, where the value of x is language specific] much as we would recite the numbers 'one' through 'ten', making it clear through the counting procedure itself that each member of the series is agreed upon as a ten-multiple of the immediately lower member. Though skepticism on the part of those encountering these systems is perhaps understandable, any a priori rejection of the possibility that such systems could exist in a traditional society is in our opinion unwarranted.

In Section 1. of this paper, we present some background information on the numeral systems of MC languages, with particular attention paid to the status of the category countable base (numeral classifier). Section 2. considers the ten-power bases of MC languages, with some attempt at reconstruction of these bases for earlier periods in the history of these languages (Section 2.1.) and a somewhat more detailed discussion of MC forms for 'ten' (Section 2.2.). In

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics,* 61-79.
*Pacific Linguistics,* C-80, 1984.

Section 3., we consider very briefly the ten-power bases of other Oceanic (OC) languages. Section 4. provides a speculative first attempt at an account of the evolution of the ten-power base series.

## 1. MICRONESIAN NUMERAL SYSTEMS

This section gives a brief review of those features of the grammar of numerals in MC languages that have some bearing on the interpretation of the ten-power base series found in many of these languages.

All MC languages reflect the Proto-Oceanic (POC) numbers 'two' through 'nine' as reconstructed in Pawley (n.d.),[4] with the sole exceptions of Marshallese (MRS), which has replaced the numbers 'six' through 'eight' by other forms, and Trukese (TRK), which has similarly replaced the numbers 'three' and 'four' in unit counting, although not in serial or higher-order counting. (These language-specific developments are not directly relevant here, however.) Except in the serial (enumerative) counting systems of some languages, these numbers always appear as first component of a bimorphemic numeral, the second component of which is a *numeral classifier*. All MC languages, again with the exception of MRS,[5] have a numeral classification system, though the number of classifiers varies from language to language. A minimal binary classification system is found in Kosraean (KSR) (Lee 1975). At the other extreme, TRK (Sugita in preparation, Benton 1968) and Kiribati (Gilbertese) (KIR) (Harrison in preparation, Trussel 1979) have some ninety classifiers, not all of which however, are in common use. Rehg (1981) lists twenty-nine classifiers for Ponapean (PNP). Sohn (1975), in a "non-exhaustive list", gives thirty-eight Woleaiean (WOL) numeral classifiers. The following table gives the numerals 'two' through 'nine' in the general counting system (with reflexes of the classifier Proto-Micronesian (PMC) *-ua < POC *pua *fruit*) and in the animate counting system (with reflexes of the classifier PMC *-manu[6] < POC *manu(k) *bird, creature*) in KIR, PNP, and WOL, representative of three of the five first-order MC subgroups.

Though we have employed the widely-used term *numeral classifier* above, it is perhaps not as appropriate for these morphemes, given their semantics, as a more neutral term such as *countable base*. Some countable bases are 'qualitative' (what Lyons 1977:460ff terms 'sortal') selected in terms of a classification of objects in the world on the basis of salient features of the inherent semantics of the objects being counted. For example:

| | | |
|---|---|---|
| KIR | uoua te boki<br>*two-general art book* | *two books* |
| | uoman ataei<br>*two*-animate *child* | *two children* |
| | uakai te nii<br>*two-plant art coconut* | *two coconut trees* |
| | uaai te tikareti<br>*two-long art cigarette* | *two cigarettes* |

Other countable bases do not reflect a hyponymic classification in this sense, but are set labels for individual classes of objects (what Benton 1968 terms 'repeaters'):

| | KIR | | PNP | | WOL | | PMC | |
|---|---|---|---|---|---|---|---|---|
| | General | Animate | General | Animate | General | Animate | General | Animate |
| *two-* | uoua | uoman | riau | riemen | riuwauw | riuwemal | *ruwa-ua | *ruwa-manu |
| *three-* | teniua | teniman | siluh | silimen | seliuw | selimel | *telú-ua | *telú-manu |
| *four-* | aua | aman | pahieu | pahmen | faauw | faamal | *fa(a)-ua | *fa(a)-manu |
| *five-* | nimaua | niiman | limau | limmen | limauw | limmal | *lima-ua | *lima-manu |
| *six-* | onoua | onomaun | weneu | wenemen | wolouw | wolomal | *ono-ua | *ono-manu |
| *seven-* | itiua | itimaun | isuh | isimen | fisiuw | fisimel | *fitú-ua | *fitú-manu |
| *eight-* | waniua | waniman | waluh | welimen | waliuw | walimel | *walú-ua | *walú-manu |
| *nine-* | ruaiua | ruaman | duwau | duwemen | tiweuw | tiwemal | *s'iwa-ua | *s'iwa-manu |

| | | | | |
|---|---|---|---|---|
| TRK | rúwéchamw<br>*two-head* | chamwen<br>*head-of* | iik<br>*fish* | *two fish heads* |
| | rúwéché<br>*two-flat object* | chéén<br>*leaf-of* | wuuch<br>*banana* | *two banana leaves* |
| | rúwéféw<br>*two-round object* | faaw<br>*rock* | | *two rocks* |
| | rúúfutuk<br>*two-meat* | futuken<br>*meat-of* | kkow<br>*cow* | *two lumps of beef* |

However, in most MC languages, a large proportion of all countable bases are 'quantitative' (what Lyons terms 'mensural'), referring to enumerable (measurable) quanta of time or space, to containers, sets or agglomerations, or to parts of a whole. For example:

| | | | |
|---|---|---|---|
| KIR | uabong | *two days* | (bong *day*) |
| | uangaa | *two fathoms* | (ngaa *fathom*) |
| | uamwangko | *two cups* | (mwangko *cup*) |
| | uarinan | *two rows* | (rinan *row*) |
| | uaatao | *two layers* | (atao *layer*) |
| | uamwakoro | *two pieces* | (mwakoro *piece*) |

The category *quantitative countable base* in MC languages also includes bases with a fixed numerical value, the ten-power bases that are the focus of this article. The morphology and syntax of these items is, in most languages, not distinct from those of other countable bases. This observation will prove significant in our account of the evolution of the ten-power base series (see Section 4.). Some examples are:

| | | | | |
|---|---|---|---|---|
| KIR | uangaun | *twenty*<br>*two-unit of ten* | (ngaun | *unit of ten*) |
| | uabubua | *two hundred* | (bubua | *unit of hundred*) |
| | uangaa | *two thousand* | (ngaa | *unit of thousand*) |

## 1.1. Serial (enumerative) counting and the number *one*

With the exception of MRS, all MC languages (and many other OC languages as well)[7] possess a serial or enumerative counting system, which, in MC languages, does not involve countable bases (classifiers). The numbers in this system are not used as nominal adjuncts; they are used either in abstract counting or in enumerating a series. The following table presents the serial counting systems from *one* to *ten* of representative MC languages.[8]

| Serial/enumerative counting | | | | | | | |
|---|---|---|---|---|---|---|---|
| | KIR[9] | KSR | PNP | MOK | TRK | WOL | PMP |
| *one* | (te)era | sra | ehd | oahd/ehd | eet | yet [yeetA] | *-sa |
| *two* | uua | lo | (a)ri | (a)ri | (té)rúú | riuw [rúúwA] | *rua |
| *three* | teen | tol | (e)sil | (e)jil | één | yel [yeelI] | *telú |
| *four* | aanga | ahng | (e)peng | (oa)poang | fáán | fang [faangI] | *fangi/*fanga |
| *five* | niima | luhm | (a)lim | (a)lim | niim | lim [liimA] | *lima |
| *six* | | on | (o)un | (o)hn | woon | wol [wool0] | *ono |
| *seven* | | it | (e)is | (e)ij | fúús | fis [fiisI] | *fitú |
| *eight* | | oal | (e)wel | (a)wal | waan | wal [waalI] | *walú |
| *nine* | | yuh | (a)du | (a)du | tiiw | tiw [tiiwA] | *s'iwa |
| *ten*[10] | ikoa | singuhul | eisek | eijek | engoon | seig | |

With the exception of the serial counter for *four*, which apparently reflects a PMC *fangi (KIR aanga suggests an earlier *fanga), the serial counting forms for *two* through *nine* in all the languages reflect established POC reconstructions for the cardinal numbers. The forms for *one* appear to reflect POC *(n)sa (Pawley, n.d.) — with all the languages except KSR showing clear evidence of a prefixed increment. Pawley has also reconstructed POC *ta(n)sa *one*..., but none of the MC prefixes is a clear reflex of *ta-. If, however, the prefix on the forms for *one* were reconstructed as *te-, as suggested by KIR teera,[11] all forms could be accounted for. Moreover, there is substantial evidence, in the Solomons at least, that such a form existed outside MC as well.

A second prefix must also be reconstructed to account for the forms for *two* through *nine* in the PP languages PNP and Mokilese (MOK). This prefix, which apparently is not attested outside PP, can be reconstructed as PPP *a-.[12] The data also suggest the need for yet a third prefix to be reconstructed to account for TRK térúú *two*, but it is our belief that the té- reflects the final *sa of *te-sa *one* copied onto *ruwa *two*: yeeté-rúú. (In TRK, final vowels are deleted at the ends of words, but not before enclitics or suffixes. We suggest that *te-sa-ruwa *one-two* was treated in TRK as a single 'phrase', before final vowel deletion applied.)

The forms for *one* with countable bases involve a prefix that in most languages is also a reflex of PMC *te-:

| | | |
|---|---|---|
| KIR | teuana[13] | *one* (general object) |
| | temanna | *one* (animate creature) |
| MRS | juon | *one* |
| TRK | eew | *one* (general object) |
| | emen | *one* (animate creature) |
| WOL | seuw | *one* (general object) |
| | semal | *one* (animate creature) |

However, although PP forms for *one* also involve a prefixed countable base, the prefix is more likely to be a reflex of PPP *a- than of PMC *te-, unless one assumes irregular assimilatory lowering in forms like MOK apas *one* (long object) (cf. Harrison 1976, Rehg this volume):

| | | |
|---|---|---|
| MOK | ew | *one* (general object) |
| | emen | *one* (animate creature) |
| | apas | *one* (long object) |
| | ekij | *one* (piece) |

One must, then, reconstruct the following forms for *one*:

i) PMC *-sa *one* (serial counting), which is a reflex of POC *(n)sa *one*.

ii) PMC *te- *one*, which is affixed in KIR and the Trukic (TK) languages to *-sa in serial counting, and in MRS, KIR, and TK to countable bases. It is likely that *te is also related to the KIR common article te and to a demonstrative prefix in TRK (see Bender 1981 and Harrison in preparation). Similar forms, though usually ta, are attested in the New Hebrides and the south-east Solomons. The isomorphism between the KIR article and the number *one* is unlikely to be an accident, though it is difficult to determine which function, if either, is historically prior.

iii) PPP *a- 'unit and serial prefix', which may, however, also prove to be related to the a- prefix appearing with demonstratives in KIR (aei *this*, anne *that*, aarei *that*).

## 2. TEN-POWER COUNTABLE BASES IN MICRONESIAN LANGUAGES

As already mentioned, morphemes for *ten*, *hundred*, etc. in MC languages are countable bases that combine with the number prefixes *one* through *nine* to form numerals.[14] For example, with the countable base *ten*, we find:

| | KIR (-bwii *ten*) | PNP (-isek *ten*) | WOL (-ig *ten*) |
|---|---|---|---|
| *ten* | tebwiina | eisek | seig |
| *twenty* | uabwii | rieisek | riuweig |
| *thirty* | tenibwii | silihsek | seliig |
| *forty* | abwii | pahisek | faaig |
| *fifty* | nimabwii | limeisek | likeig |
| *sixty* | onobwii | weneisek | woloig |
| *seventy* | itibwii | isihsek | fisiig |
| *eighty* | wanibwii | welihsek | waliig |
| *ninety* | ruabwii | duweisek | tiweig |

Higher numbers such as, for example, *four hundred and thirty five* are formed in a left-to-right sequence beginning with the highest appropriate power of ten. In some languages the numerals are conjoined, while in others they are juxtaposed: Thus:

KIR abubua ao tenibwii ma nimaua *435*
*4-100 and 3-10 and 5-general*

MOK pahpwiki jilihjek limoaw *435*
*4-100 3-10 5-general*

WOL faabiugiuw me seliig me limauw *435*
*4-100 and 3-10 and 5-general*

Unlike Indo-European ten-power numerals, which, except for relatively recent forms like 'million', do not exceed $10^3$ (*thousand*), the ten-power bases of some MC languages go as high as $10^9$. The following chart presents the unit numeral for the ten-power morphemes (i.e., 1 x $10^2$, etc.) in a representative sample of MC languages:[15]

| KIR | MRS | KSR | PNP | TRK | WOL | CRL | |
|---|---|---|---|---|---|---|---|
| tebwiina | jo͞noul | singuhul | eisek | engoon | seig | seigh | $10^1$ |
| tebubua | jibukwi | siofok[16] | epwiki | epwúkú | sebiugiuw | ebwúghúw | $10^2$ |
| tengaa | jerapen[17] | | kid | engérέw | sangeras | sangaras | $10^3$ |
| terebu | | | nen | ekit | sen [sennA] | ssel | $10^4$ |
| tekuri | | | lopw | | selob | | $10^5$ |
| teea | | | rar | | sepiy | | $10^6$ |
| tetano | | | dep | | sengit | | $10^7$ |
| tetoki | | | sapw | | sangerai | | $10^8$ |
| | | | lik | | | | $10^9$ |

Although the extent to which the entire set is known by all speakers of a language seems to vary, in no case is it difficult to find speakers willing to volunteer the full series. Furthermore, as mentioned earlier, speakers are in total agreement as to the values of the numbers.

In some languages, numbers above $10^3$ are more usually formed on the English pattern (often with Eng. borrowings like KIR mirion *million* and birion *billion*. For example:

| | | |
|---|---|---|
| KIR | teningaa | *3,000* |
| | tenibwii tengaa | *30,000* |
| | tenibubua tengaa | *300,000* |
| | teniua te mirion | *3,000,000* |
| | tenibubua ao tenibwii tengaa | *330,000* |
| | tenibubua ao tenibwii ma teniua tengaa | *333,000* |

In the traditional KIR system, '300,000' would be rendered

| | | | | |
|---|---|---|---|---|
| KIR | tenikuri | ao | tenirebu | ao teningaa |
| | *3-$10^5$* | *and* | *3-$10^4$* | *and 3-$10^3$* |

## 2.1. Reconstructing PMC ten-power bases

Ignoring the base 'ten' itself (see Section 2.2.), it is possible to reconstruct five of the ten-power base series on the basis of cognates found in at least two lower-order subgroups of MC (ULI = Ulithian, PUL = Puluwat):

| | | | | | |
|---|---|---|---|---|---|
| PMC | *pukua | *k(u,i)si[18] | *lopwa | *sep(u,i) | *nena |
| KIR | bubua | kuri | | rebu | |
| MRS | bukwi | ??jo-kde-n | | ??dep | |
| KSP | foko | | | | |
| PNP | pwiki | kid | lopw | dep | nen |
| TRK | pwúkú | (k)kit | | | |
| CRL | bwúghúw | | | | sse-l |
| ULI | buxuy | | | | |
| WOL | biugiuw | | lob[lobᵂA] | | se-n [sennA] |
| PUL | pwúkúw | kkit | | | |

PMC *púkua *hundred* is reflected throughout MC. It is problematic only in KIR, where one must assume an irregular change *k > b if the KIR form is in fact cognate. PMC *k(u,i)si is also widely distributed, though its value is not consistent ($10^5$ in KIR, $10^3$ in PP, and $10^4$ in TRK and PUL). MRS jokden *ten pairs* is doubtful, since MRS d is not a regular reflex of PMC *s. PMC *lop'a is reflected only in PP and WOL. PMC *sep(u,i) is reflected in KIR as $10^4$ and in PP as $10^7$.

The reconstruction PMC *nena requires some comment. It is reflected in TK and PP with the same value ($10^4$). In the former, however, the reflexes at first glance might appear to be less than secure. For WOL sen [sennA], we hypothesise loss of *e between identical consonants (i.e., *te-nena > se+nnA), which is a regular development in TK. This same process is assumed in the development of CRL ssel. In this form, however, the mora count has been redistributed, so that it is now the initial consonant that is geminate (where a geminate consonant counts as a mora): *te-nena > se+nna > ssena > ssel. This type of development, too, is widespread in Trukic.

Two additional reconstructions can be made for PTK:

| | | |
|---|---|---|
| PTK | *t''e-ngarat''u | *(p)piya |
| TRK | e-ngérew | |
| CRL | sa-ngaras | ppiy |
| ULI | sa-ngaras | ppiy |
| WOL | sa-ngeras | piy |
| PUL | ye-ngeráy | |

PTK *(p)piya is in fact a noun meaning *'sand'* (see Section 4.). PTK *t'e-ngarat'u *10*$^3$ has no cognates elsewhere in MC but is widely reflected elsewhere. It is clearly a reflex of PAN *Ratus, POC *Ratu *hundred* (Blust 1972), with the unit prefix PMC *te and reflex of the numerical ligature POC *nga (see Pawley, n.d.). KIR -ngaa *10*$^3$ is a possible cognate for PTK *ngarat'u *10*$^3$ but only if one assumes a very irregular loss of the historical final syllable. (The expected KIR form would be *ngaati).

It is significant that WOL shows a contrast between sangaras *1,000* and sangerai *100,000,000*. The former, as already discussed, is a reflex of POC *nga-Ratu *100*, while the latter is undoubtedly related to PPN *lau *countless, indefinite number'* (Biggs, et al. 1970), which is taken by Pawley (n.d.) to reflect his POC *dau *hundred, unit of hundreds*. So far as we are aware, WOL is unique in independently reflecting both of these POC reconstructions. The failure of the other TK languages to reflect both forms, however, may be the result of recent phonological developments which would result in *Ratu and *dau being reflected identically: TRK and Mortlockese (MRT) engérew and PUL yengeráy are as likely to be reflexes of one as the other POC reconstruction. (As Codrington (1885:249) noted, the isomorphism between the word for 'hundred' and that for 'leaf' [POC *dau] is probably not accidental. Note, however, that all MC languages reflect a POC nasal grade *ndau *leaf* but oral grade *dau *hundred* where this item is reflected.)

In KIR, three (four if KIR ngaa has TK cognates) of the ten-power bases above 'ten' (10$^2$ through 10$^5$) have cognates in other MC languages. Of the remaining three, tano and toki have nominal interpretations in KIR, the former meaning *sand, soil, ground* and the latter *end, limit*. In PP, sapw *10*$^8$ also means *land* and lik *10*$^9$ means *outside*. PNP rar *10*$^6$, as a noun, has the meaning *finger coral*. It is possible that this latter isomorphism is not accidental, given the nature of 'finger coral', an object that is frequently found in small fragments on the shore. MOK evidence with regard to this item is problematic, however, since *finger coral* is MOK lar while *10*$^6$ is MOK rar. All the TK forms have numeral cognates, either intra-MC or extra-MC, except PTK *(p)piya *sand*, WOL -ngit *10*$^7$ and CRL -púngút *10*$^5$. (It is possible, however, that these latter forms are related in some way to Uripiv ongut *100* and similar forms in the northern New Hebrides (Ray 1926).)

It is interesting to note that in the case of those higher power numerals with independent nominal interpretations, although the forms themselves are not cognate in a strict sense, they appear to reflect similar semantics. Thus the highest ten-power bases of KIR and PP (KIR toki *end*, PNP lik *outside*) both carry the sense of 'limit, extreme'. In these same languages the immediate lower ten-power base (KIR tano *sand, soil*, PNP sapw *land*) also have similar nominal interpretations, as does PTK *(p)piya *sand*, the highest power base in CRL and ULI and a relatively high one (10$^6$) in WOL. These observations suggest that a similar semantic may govern the ten-power base systems of all MC languages, even in cases where the forms themselves are not cognate.

## 2.2. Micronesian forms for 'ten'

The following forms for 'ten', or for groups of ten, are extant in MC languages (PUA = Pulo Annian):

| | | | | |
|---|---|---|---|---|
| KIR | tengaun | tebwiina | ?ikoa | |
| MRS | joṉoul | | | |
| KSR | singuhul[19] | | | |
| PNP | ngoul | | ehk | eisek |
| TRK | engoon[20] | | -ik | |
| PUL | yengool | | heeyik | |
| CRL | | ääf | seeigh | |
| WOL | sengaul | seyaf | seig | |
| ULI | | | seyex | |
| PUA | | | deikI | |
| MRT | yengool | | sɛɛk | |

Reflexes of PMC *ngaulu *ten* are found in all subgroups of MC, and serve as the sole forms for 'ten' in MRS and KSR. The KIR reflex is used in all counting systems except the general (-ua) system (in which case KIR tebwiina *ten* is used). The PNP reflex is restricted to counting days, food prepared in an earth oven, multi-stemmed plants, and small pieces.[21] WOL sengaul has the restricted interpretation *ten groups*. Cognates of PMC *ngaulu *ten* are widely distributed outside MC as the number 'ten'. Its base is reconstructed as POC *pulu *ten*, usually reflected with a unit prefix and the ligature *nga, parallel to POC *Ratu and *dau.

CRL ääf *ten* does not appear to be current, though reported in Fritz (1911). WOL seyaf means *ten pieces of copra* and is also used in counting coins and valuable shells. The KIR root -bwii *ten* may be related to MOK (e)-pwi *some, several* and to PP *pwihn and TK pwii *group* (with PP languages showing a reflex of the suffix *-na). (This latter form has been reconstructed as PMC *p'utu(-na) *group, flock, school*, largely on the basis of MRS bwijin *group* (PMC *t > MRS j). Though loss of PMC *t is not unexpected in PP and some TK languages in this environment, it would be unusual for KIR.)[22]

The history of the remaining forms for 'ten' is complex. For the TK forms and PNP ehk we reconstruct a base *-ke *ten*. The PNP form derives from this base, with a unit prefix (probably *a-) by regular historical changes (see further below). The TK forms require us to postulate an optional prefix i-[23] (PTK *t''e-i-ke *ten*) whose source and function is unclear. However, these same elements, in a different order, also yield PNP eisek *ten* (< PPP *a-i-te-ke).

Although TRK, PUL, and MRT do not provide direct evidence for *-ke in the forms for *twenty* (rúwe), *fifty* (lime), *sixty* (wone), and *ninety* (ttiwe) (i.e. when the number morphemes end in a non-high vowel), the loss of *k in this environment is not unexpected: e.g. *ruwa-ke *twenty* > *rúwa-e > rúwee > rúwe. After historical high vowels, PNP -ehk has the alternate -akan (PNP siliakan *thirty*, isiakan *seventy*, weliakan *eighty*), the shape of which is problematic. The final -n may be a reflex of the *-na suffix, but why it should appear in these forms is unclear, since it is usually restricted to units (*one, ten*, etc.) rather than in multiples. The low vowel of -ka- is likewise unexpected. KIR ikoa *ten (pair counting)* is only questionably a reflex of an earlier *i-ke, since no account can be given for the final a of the KIR form.

The reconstructed *-ke *ten* appears to have extra-MC cognates, but with the value *one*; for example, the form ke *one* is widespread in the south-east Solomons. (Note also Roviana (ROV) manege *ten*.) There may be some relationship between these forms and the reconstructed POC *ta-(n)sa-kai *one* (Pawley n.d.), with a unit prefix, though the *-ke ∿ *-kai alternation is perhaps problematic. One might also note, within MC, KSR soko, a possible reflex of POC *(n)sa-kai, which becomes more plausible if one can assume a pre-KSR *sa-kau *one*.

## 3. TEN-POWER BASES OUTSIDE MICRONESIAN

Ten-power base systems of the sort described here, though perhaps most developed in MC, are attested in many other OC languages.[24] In PN, for example, though the series ends in most languages at $10^3$ ('thousand'), Tongan (TON) has numerals up to $10^5$. The most extensive ten-power base system thus far encountered is, in fact, that of NUK, which reaches $10^{10}$. Other OC languages, especially in the Solomons (e.g. Bugotu (BUG), and Kia (KIA) of Guadalcanal), Motu (MTU), Nguna (NGU), and ROV, have ten-power bases above $10^3$. Thus (SAM = Samoan):

| TON | SAM | NUK | MAO | BUG | KIA | MTU | NGU | ROV | |
|---|---|---|---|---|---|---|---|---|---|
| hongofulu | sefulu | hulu/mada | ngahuru | sa-lage/ hangavulu | tazo/ (ŋa)fulu | gwauta | (dualima) | manege/ -ŋavulu | $10^1$ |
| teau/ngeau | selau | lau | rau | hathangatu | gobi | sinahu | põnotia | gogoto | $10^2$ |
| afe | afe | mano | mano | toga | toga | dahu | manu | tina | $10^3$ |
| mano | | semada | | mola | vuro | gerebu | tivilia | vuro | $10^4$ |
| kilu | | seguli | | feferi | feferi | domaga | | | $10^5$ |
| | | seloo | | vuthea | tahozea | | | | $10^6$ |
| | | sengaa | | vathegila | | | | | $10^7$ |
| | | semuna | | | | | | | $10^8$ |
| | | sebugi | | | | | | | $10^9$ |
| | | sebaga | | | | | | | $10^{10}$ |

We will not attempt any reconstructions from these data, except to point out that most of the forms for 'ten' reflect the POC base *pulu. The bases POC *dau and *Ratu, considered in Section 2.1., are also reflected in some of the forms above. (For PPN reconstructions up to $10^3$, see Biggs et al. 1970; Pawley (n.d.) gives some reconstructions for POC.)

## 4. EVOLUTION OF THE TEN-POWER BASE SYSTEM

Though extensive ten-power base systems are extant in a number of widely separated areas of Oceania, it is unlikely that forms above $10^3$ (or possibly $10^2$) can be reconstructed for POC (or for PMC). One observes, for example, that:

i) few of the higher ten-power bases (above $10^3$) are reconstructable beyond the very lowest order subgroups.

ii) in many instances, what appear to be cognate items have different ten-power values in different languages.

iii) in some languages, the morphology of ten-power bases above $10^3$ is distinct from that of the lower ten-power bases. Thus PP ten-power bases above $10^3$ do not take the unit prefix e-, while in NUK those below $10^4$ do not take the prefix se-.

These observations suggest that the higher ten-power bases have to some degree a history distinct from that of the lower. In our view, they may have developed as numbers at a more recent historical period. Similarly, the ligature *nga is found only (but not always) with the bases *pulu *ten*, *Ratu *hundred*, *dau *hundred, large number* (where these bases are reflected in extant languages), suggesting an older numeral morphology fossilised in some reflexes of the oldest ten-power bases. Though it is possible that an older extensive ten-power base system may have broken down and subsequently been reintroduced with new items, it is more likely that the systems found are the result of a number of independent innovations. That such systems developed is not in our opinion completely accidental, as we will attempt to demonstrate below.[25]

As stated above, evidence from MC (and throughout OC) suggests that the ten-power bases, even those of the simplest systems, were morphosyntactically not numbers of the same sort as the numbers 'one' through 'nine' but were countable bases (quantitative classifiers) that were themselves counted. That is, the ten-power bases can best be interpreted as 'unit of $10^x$' and counted as 'one unit of $10^x$', 'two units of $10^x$', etc. The ten-power bases themselves are, then, countable bases distinct from other countable bases only in that their sole interpretation is a mathematical abstraction,[26] a fixed quantity, and in that they form a semantic set based on another mathematical abstraction, a ten-multiple series.

However, most of the items in the ten-power base series may not always have had a numerical interpretation, but are likely to have begun as quantitative classifiers with meanings like 'pile', 'group', 'heap', or more specific meanings implying quantity, such as 'sand' or 'leaf'. There is ample evidence in the ten-power base systems of MC and other OC languages that many of these items had such interpretations; for example, PTK *(p)piya *sand*, functioning as a higher ten-power base in a number of TK languages; KIR tetano *ground, soil*, $10^7$, tetoki *end, completion*, $10^8$; PNP sapw *land*, $10^8$, lik *outside*, $10^9$; PPN *lau *leaf*, $10^2$.

The dynamic posited for the evolution of the ten-power base system is a simple one: a countable base (quantitative classifier) came to be incorporated into a mathematical series based on increasing powers of ten, the basis of which (on present evidence, as far as $10^3$) was already present in the proto-language. In some instances, the base in question may have first become associated with a specific numeral value, as possibly in the case of KIR tebwiina *ten* (MOK -pwi *several*, TK pwii- *group, school, flock*), which may have been associated with a group whose conventional extent was 'ten', and later came to represent the abstraction 'ten' alone. (Note that WOL sengaul means *ten groups (of eight or ten)*, possibly its original interpretation, or possibly the converse development to that described above.)

On the other hand, we also find it plausible to conceive of individual ten-power bases being incorporated into the existing series simply by being conventionally 'tacked on at the end', as would be the case were English to develop a word for 'ten thousand' from a sequence 'ten, hundred, thousand, a heap'. The capacity for mathematical abstraction of this sort is certainly within human potential.

The sole linguistic prerequisite for the development of ten-power base systems in the manner described above is the existence of a grammatical category 'countable base (numeral classifier)'. The larger and more open this category is, the more likely it is that new ten-power bases will be added to the series through time, though it does not follow that this development need take place. It is interesting to note that it is precisely those MC languages that have largely abandoned the numeral classifier category (MRS and KSR) that have the most limited ten-power base systems. It is not clear whether more extensive systems simply never developed in these languages, or have eroded through time along with the classifier system.

In order to account in this way for the existence of ten-power base systems outside of MC we must assume that the category 'countable base (numeral classifier)' is not a MC innovation, but existed also at earlier periods in the history of other OC languages. There is some evidence that this was in fact so. Residual sets of countable bases are a feature of many PN languages (for example, TON tekau *score*, uangakau *two score;* tetula *score (of thatch)*, uangotula *two score (of thatch);* tekumi *ten fathoms*, uangokumi *twenty fathoms*).[27] Similar quantitative countable bases are attested elsewhere — Nggela na kua *ten eggs*, na banara *ten baskets of food*, na gobi *ten canoes*, na paga *ten non-human animates*.

Qualitative countable bases, on available evidence, are not widespread in OC. Outside of MC, this category has been reported on San Cristobal and Malaita (Pawley 1972), on Bougainville (Ross 1981), and in the Trobriands. The number of qualitative classifiers reported in these languages is, however, apparently considerably smaller than in MC languages.

Extra-MC evidence for the existence of a 'countable base' category is so fragmentary that it is difficult to draw more than the most speculative conclusions regarding the level for which such a category should be reconstructed, or regarding the nature of the system (for example, whether qualitative countables evolved after quantitative ones, or whether there was an extensive system of qualitative classifiers at some earlier period in the history of OC languages that has been eroded to a greater or lesser degree everywhere but in MC). Such observations as the fact that the countable bases of TON appear with reflexes of the ligature *nga suggest that the category may have had a long history.

## NOTES

1. The term 'Micronesian' is used here in the same sense as the term 'nuclear Micronesian' of Bender (1971), including Nauruan (relevant data from which is, however, not available), Kiribati (KIR) (Gilbertese), and all languages of the United States Trust Territory of the Pacific Islands except Chamorro, Yapese, Palauan, Nukuoro (NUK), and Kapingamarangi. As a working hypothesis we will assume the existence of a Micronesian (MC) subgroup of Oceanic (OC), much as described in Marck (1977), though the integrity of such a subgroup has yet to be demonstrated conclusively. Within MC, we assume the existence of a Trukic (TK) subgroup (see Jackson pp. 259-280 in this volume), and a Ponapeic (PP) subgroup, including Ponapean (PNP), Ngatikese, Mokilese (MOK), and Pingelapese. We take the liberty here of making PMC reconstructions on the basis of cognates between any two first-order subgroups of MC, in full recognition of the fact that this procedure may be methodologically questionable.

2. In citing examples, the orthographies used in the standard reference works (see References) are employed. Where deemed necessary, phonetic detail is given in square brackets. (Data cited from languages for which no reference works are available are given in the transcription of the source.)

3. Cognates of the first four HAW numbers have base-ten interpretations; for example, Maori (MAO) rau *100*, mano *1000*, Fijian (FIJ) tini *10*, Motu (MTU) ge-rebu *10,000*. That 'four' is a common factor in the interpretations of the HAW forms is related to the particular status of 'four' in Hawaiian culture. HAW forms for 'hundred' and 'thousand' are Eng. borrowings, hanele and kaukani, respectively.

4. The number 'one' is historically problematic — see Section 1.1.

5. MRS numerals 'one' through 'three' appear to reflect the Proto-Micronesian (PMC) general classifier *-ua: juon *one* (see Section 1.1.), ruo *two*, jilu *three*, albeit somewhat opaquely. MRS emān *four* appears to reflect the PMC animate classifier *-manu. We can thus assume that earlier stages of MRS had a numeral classifier system.

6. Other classifiers are also reconstructable for PMC.

7. Codrington (1885) notes the existence of serial counting systems in several languages of the Solomons. Such systems are also found in Roviana (ROV) and Rotuman.

8. In MC languages, once 'ten' has been reached in serial counting, the series begins again at 'one'. The same procedure is followed once 'twenty' has been reached, etc.

9. The KIR forms are used in pair counting.

10. The forms for 'ten' will be considered in Section 2.2.

11. The long vowel of KIR teera, as well as the long vowels in other KIR, PNP, MOK, TRK, and WOL forms is the result of a regular phonological process. See Rehg (this volume).

12. MOK ehd *one* almost certainly reflects *te-sa; oahd, on the other hand, may reflect PPP *a-sa; that is, the *a- prefix found before number roots 'two' through 'nine' in PNP and MOK may have been extended to this form for 'one'. It is noteworthy that MOK ehd is used when counting out the ten-power series (*one-ten-hundred-thousand...*), while oahd is used only in unit counting (*one-two-three...*).

13. The suffix -na accompanies all unit value countable bases in KIR except those of the ten-power series. (KIR tebwiina *ten* takes -na except when a conjoined numeral follows: for example, KIR tebwii ma teuana *eleven*). MRS juon *one* is cognate with KIR teuana *one* (see Bender 1981).

14. Countable bases above 'hundred' in PP (as well as PNP ngoul *ten*) do not take the unit prefix e-, although they do take other number prefixes: PNP epwiki *one hundred*, riepwiki *two hundred*, but kid *one thousand*, riekid *two thousand*.

15. All the languages represented here, except MRS and KSR, have other bases with the value 'ten' — see Section 2.2. CRL stands for Saipan Carolinian.

16. KSR -foko *hundred* appears after the number prefixes 'two' through 'nine'.

17. MRS -rap(e)- *thousand* is archaic, now largely replaced by the Eng. loan toujin. Note the form MRS limādep *five thousand*.

18. The correspondences between KIR -kuri $10^5$ and the PP and TK forms are not problematic. It should be pointed out, however, that the KIR form might be a PN loan, corresponding to NUK seguli $10^5$ and TON kilu $10^5$. Goodenough and Sugita (1980) give TRK e-kit $10^4$. Our checking of this form with a native speaker revealed a geminate /kk/, as in PUL -kkit.

19. KSR -nguhul *ten* has the variants -ngoul and -ngaul after some number prefixes.

20. TRK -ngoon *ten* is not used for multiples of ten ('twenty', 'thirty', etc.); TRK -ik (∿ -e — see below) appears in ten-multiples with the prefixes 'two' through 'nine'.

21. PNP koadoangoul, a serial counting *ten* or *hundred*, reflects the same base, with a fossilised causative prefix (PMC *ka-) and a reflex of the unit number PMC *sa (> PNP doa).

22. WOL also fails to show a reflex of PMC *t in this form, which suggests the reconstruction PMC *p'u'(t)u' (-na); that is, it is MRS, not KIR or WOL, that is aberrant.

23. Neither ULI nor MRT shows evidence of the prefix *i-. There is also evidence (see below) that the prefix is not reflected in all TRK or PUL forms for 'ten'.

24. Extended ten-power base systems may be found in Western Austronesian languages, as well. Although we have not been able to systematically check the data for such languages, we find it very interesting that two sources have provided the following Ilokano forms: sangapulo *ten*, sangagasut *one hundred*, sangaribu *one thousand*, sangalaksa *ten thousand*, riwriw *one million*. (Constantino, 1971)

25. Girschner (1906), after describing the PNP ten-power base system, comments that 'expressions for still higher numbers have been introduced by missionaries but are not current with the natives'. If Girschner's remarks are in fact directed towards the forms in question here, it should be obvious that there

is no evidence for such an origin for the ten-power bases. We are grateful to Ken Rehg for bringing this reference to our attention.

26. It should be pointed out again, however, that some MC languages also show a 'qualitative' distinction in forms for 'ten' (see Section 2.). In addition to those cases already cited we might note MRT yengool *ten* (inanimate) and seek *ten* (animate). Similar contrasts are apparently found outside of MC, particularly in the Solomons; for example, Nggela na gobi *ten canoes,* na paga *ten puddings, ten pigs, ten birds, ten fish, etc.,* na pigu *ten coconuts, ten breadfruit, ten crabs, ten shellfish,* na banara *ten baskets of food,* na mola *ten baskets of nuts,* na gaibata *ten bunches of bananas,* as compared with e hangavulu *ten (counting)* (Codrington, 1885). Such contrasts are not found with bases above 'ten'. In MC at least, the primary meaning of the bases in question remains quantitative, in that the MC forms do not have *specific* lexical interpretations, in apparent contrast to at least some of the forms Codrington cites.

27. PPN *kumi *ten fathoms* is reflected as HAW 'umi *ten*.

## BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

1981 A fossilized article in Marshallese. In Jim Hollyman and Andrew Pawley, eds *Studies in Pacific languages & cultures: in honour of Bruce Biggs,* 209-228. Auckland: The Linguistic Society of New Zealand.

BENTON, Richard A.

1968 Numeral and attributive classifiers in Trukese. *Oceanic Linguistics* 7/2:104-146.

BIGGS, Bruce, D.S. WALSH, and Jocelyn WAQA

1970 *Proto-Polynesian reconstructions with English to Proto-Polynesian finder list.* Interim listing January 1970. Working Papers in Linguistics, Department of Anthropology, University of Auckland.

CODRINGTON, Robert H.

1885 *The Melanesian languages.* Oxford: Clarendon Press.

CONSTANTINO, Ernesto

1971 *Ilokano dictionary.* PALI Language Texts. Honolulu: University of Hawaii Press.

ELBERT, Samuel H.

1972 *Puluwat dictionary. Pacific Linguistics,* C-24.

ELBERT, Samuel H. and Mary Kawena PUKUI

1979 *Hawaiian grammar.* Honolulu: University Press of Hawaii.

FRITZ, Georg

1911 *Die Zentralkarolinische Sprache: Grammatik, Übungen und Wörterbuch der Mundart der westlich von Truk liegenden Atolle, insbesondere der Saipan-Karoliner.* Berlin: G. Reimer.

GIRSCHNER, Regimental Physician Dr.

1906 Grammatik der Ponapesprache. *Mitteilungen der Seminar für Orientalische Sprachen: I Abteilung,* 73-126. Berlin. (Translated by S. Tully, 1975, as 'Ponapean grammar'. Mimeo.)

GOODENOUGH, Ward H. and Hiroshi SUGITA

1980 *Trukese-English dictionary.* Memoir 141 of the American Philosophical Society. Philadelphia: American Philosophical Society.

HARRISON, Sheldon P.

1976 *Mokilese reference grammar.* Honolulu: University Press of Hawaii.

in prep *A grammar of Gilbertese.*
aration

JACKSON, Frederick H.

1983 Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. See pp. 259-280 in this volume.

LEE, Kee-dong

1975 *Kusaiean reference grammar.* Honolulu: University Press of Hawaii.

LYONS, John

1977 *Semantics,* vol.2. Cambridge: Cambridge University Press.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

PAWLEY, Andrew K.

1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Readings in Oceanic culture history,* vol.3:1-142, Pacific Anthropological Records 13. Honolulu: Bernice Pauahi Bishop Museum.

n.d. A Proto-Oceanic grammar. MS.

RAY, Sidney H.

1926 *A comparative study of the Melanesian languages.* Cambridge: Cambridge University Press.

REHG, Kenneth L.

1973 On the history of Ponapean phonology. ***Working Papers in Linguistics, University of Hawaii*** 5/8:17-55. See also Rehg 1983.

1979 The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See also Rehg 1983.

1981 ***Ponapean reference grammar***. Honolulu: University Press of Hawaii.

1983a On the history of Ponapean phonology. See pp. 281-316 in this volume.

1983b The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See pp. 52-58 in this volume.

ROSS, Malcolm

1981 Proto-Oceanic from the top down. MS. University of Papua New Guinea.

SOHN, Ho-min

1975 ***Woleaian reference grammar***. Honolulu: University Press of Hawaii.

SUGITA, Hiroshi

in prep-aration Trukese reference grammar.

TRUSSEL, Stephen

1979 ***Kiribati (Gilbertese) grammar handbook***. Peace Corps Language Handbook. Brattleboro, Vermont: The Experiment in International Living.

Harrison, S.P. and Jackson, F.H. "Higher numerals in several Micronesian languages". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:61-79**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.61*

# THE IMPACT OF BORROWING ON PALAUAN

Lewis S. Josephs

## 1. INTRODUCTION

Within the last century alone, the Palau Islands of the Western Carolines have been under the influence of four different foreign administrations. Spanish culture, gradually introduced into Palau after the annexation of the islands by Spain in 1686,[1] lasted until the end of the Spanish-American War. In 1899 Spain sold the islands, together with other possessions in the Carolines group, to Germany, which administered them briefly until the armistice terminating World War I. After World War I, the League of Nations granted Japan a mandate to the Carolines, including Palau. In spite of its withdrawal from the League in 1935, Japan did not relinquish the mandated islands, but instead put them to its own military and economic use. Palau's principal town, Koror, was made the administrative headquarters of all the Caroline Islands, and the islands of Peleliu and Angaur, south-west of Koror, were heavily fortified. Japan's defeat in World War II marked the end of its control over Palau and the Carolines, and on July 19, 1947, the United States was granted United Nations trusteeship over the area (henceforth called the Trust Territory of the Pacific), which is now in the process of termination.

The impact of four different non-indigenous languages — three European and one Asiatic — upon the language of Palau has been understandably great. Large numbers of lexical items borrowed from Spanish, German, Japanese, and English are used frequently in everyday speech. As we might expect, the majority of these represent names for items or concepts originally alien to Palauan culture, but it is clear that quite a few foreign borrowings have also replaced, or at least come into competition with, native lexical items. Many borrowed words have been 'Palauanised' to some degree or other — that is, they have been subjected to the complex derivational and inflectional processes of Palauan noun and verb morphology and have been introduced into certain characteristic syntactic constructions. Although many foreign lexical items have been taken into Palauan without any change in meaning, a significant number show a rather wide variety of semantic modification and distortion. Judging from a purely statistical count of foreign borrowings and from the over-all effect of Palauanisation, the influence of Japanese has obviously been the most pervasive, in spite of the fact that Japan controlled Palau for less than three decades. From the same points of view, English comes in second (and will presumably continue to have a strong influence), while Spanish and German, respectively, show considerably weaker impact.

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 81-123. *Pacific Linguistics*, C-80, 1984.

## 2. PALAUAN LOAN WORD PHONOLOGY

In general, the influx of foreign words into Palauan (PAL) has had surprisingly little effect on the language's phonological system. With a few relatively minor exceptions, foreign borrowings are always modified or distorted to conform with the phonemes and phonotactics of PAL, and in many interesting cases, these borrowings undergo phonological rules just as if they were native stems. The inventory of phonemes and phoneme combinations has increased only slightly under the influence of foreign sounds: most obvious is the appearance of the fricative /h/ (pronounced [h] in words of Jp. and Eng. origin and [x] in several items from Sp.) and the affricates /ts/ (corresponding to Jp. [č] or [t͡s] and Eng. [č]) and /dz/ (corresponding to Jp. [ǰ]). Whereas chronologically older loanwords from Sp., Ger., and Jp. have been almost totally assimilated into the PAL phonological system, more recent loanwords from Eng. do not exhibit the same degree of Palauanisation. This is undoubtedly due to the fact that they represent the newest wave of foreign intruders and also to the fact that many educated speakers of the younger generation are essentially bilingual with Eng. Since these speakers have learned the correct Eng. pronunciation, they often preserve it when Eng. words are borrowed into PAL conversation.

In an attempt to illustrate the overwhelming tendency to reinterpret foreign borrowings in terms of the PAL phonological system, the sections below present a somewhat simplified sketch of PAL phonology[2] in which various classes of sounds are examined with emphasis on their occurrence in loanword vocabulary. Reference is made where necessary to the phonological systems of the contributing languages and to the relevant orthographic and (for Jp.) Romanisation systems.

### 2.1. Palauan consonants and their use in loanwords

In the chart below, the PAL consonant phonemes are presented in standard PAL orthography according to manner and point of articulation:

Chart of PAL consonants

| Manners of Articulation | Points of Articulation | | | | |
|---|---|---|---|---|---|
| | bilabial | dental | alveolar | velar | glottal |
| Stops < voiceless | | t | | k | ch |
| Stops < voiced | b | d | | | |
| Fricative | | | s | | |
| Nasals | m | | | ng | |
| Liquids | | | r, l | | |

Figure 1

As we will see below, the rather restricted consonantal system of PAL requires considerable distortion of nonindigenous consonantal sounds that have no exact PAL counterparts.

2.1.1.

There are only two *labial* consonant phonemes in PAL — namely, nasal m and non-nasal b. The bilabial nasal m poses no problem and is used in loanwords whenever the contributing language has m, as the following examples illustrate:

| (1) | | Loan Source[3] | | PAL[4] | |
|---|---|---|---|---|---|
| | Sp. | maíz | *corn* | maís | |
| | | medalla | *religious medal* | medália | |
| | Ger. | Maschine | *machine, motor* | mesíl | |
| | | Turm | *tower* | turm | *steeple* |
| | Jp. | mado | *window* | mádo | |
| | | maguro | *tuna* | manguró | |
| | Eng. | movie | | mubí | |
| | | rum | | rrom | *liquor* |

The non-nasal labial phoneme b is a bilabial stop with two major allophones. Syllable-initially before an l or a vowel, the *voiced* allophone [b] occurs (e.g. blái [bláy] *house*, bilís [bilís] *dog*, rubák [rubák$^{h}$] *old man*). Adjacent to another consonant in the same syllable (except l) and in word-final position, we find the *voiceless* allophone [p] (e.g. btuch [ptu?] *star*, tbak [tpak$^{h}$] *my spittle*, chab [?ap] *ashes*).

All labial obstruents in words borrowed from Sp., Ger., and Eng. are reinterpreted as the appropriately conditioned allophone of PAL /b/. Thus, the stops p and b and the fricatives f and v all become PAL /b/,[5] as the following indicate:

| (2) | | Loan Source | | PAL | |
|---|---|---|---|---|---|
| | Sp. | padre | *priest* | badré | |
| | | Pascua | *Easter* | Baskuá | |
| | | plato | *plate, dish* | belatóng | |
| | | purgatorio | *purgatory* | burkatório | |
| | | trompeta | *trumpet, bugle* | trombetáng | |
| | | bandera | *flag, banner* | bangderáng | |
| | | barril | *barrel* | barríll[6] | |
| | | Biblia | *Bible* | Bíblia | |
| | | botella | *bottle* | butiliáng | |
| | | diablo | *devil* | diablóng | |
| | | calabaza | *pumpkin* | kal(e)basáng | |
| | | calabozo | *dungeon, cell* | kelebús | *jail, prison* |
| | | sábado | *Saturday* | sebadóng | |
| | | cebolla(s) | *onion(s)* | sebuliás[7] | |
| | | farol | *lantern* | barór | *table lamp* |
| | | virhén | *virgin* | birhén | |
| (3) | Ger. | Papier | *paper, document* | babiér | *paper, letter, book* |
| | | Post | *post office* | bost | |
| | | Kapitän | *captain* | kabitéi | |
| | | Lampe | *lamp, lantern* | lámbei | |
| | | Schlips | *necktie* | slibs | |
| | | Brief | *letter* | beríb | |
| | | Bild | *picture* | bilt | *holy picture* |
| | | Schraube | *screw* | seráub | |

| | | | | |
|---|---|---|---|---|
| | Tafel | *blackboard* | tabér | |
| | Fenster | *window* | béngster | |
| | Grammophon | *phonograph* | karmoból | |
| | Maschinengewehr[8] | *machine-gun* | mesilkebiér | |
| | auswendig | *by heart* | chausbéngdik | *memorise* |
| (4) Eng. | pipe | | báeb | *pipe, windpipe* |
| | piston | | bistóng | |
| | present | | breséngt | |
| | aspirin | | chasberíng | |
| | hospital | | osbitár | |
| | company | | kombalíi | |
| | sheep | | sib | |
| | box | | baks | |
| | beer | | biáng | |
| | bishop | | bisób | |
| | club | | klab | |
| | table | | tébel | |
| | flour | | blauáng | *bread* |
| | office | | obís | |
| | softball | | sabtból | |
| | valve | | barb | |
| | verandah | | berangdáng | |
| | glove | | kurób | |
| | navy | | neibí | |
| | shovel | | sébel | |

As the words in (2-4) illustrate, any labial from Sp., Ger., or Eng. becomes PAL /b/ realised phonetically as the indicated allophone in the selected examples below:

(5) [b]

(i) syllable-initially, before l:

Sp. diablo = PAL diablóng [diablóŋ]

Eng. flour = PAL bláuang [bláwaŋ]

(ii) syllable-initially, before a vowel:

| | | | | |
|---|---|---|---|---|
| Sp. | Biblia | = PAL | Bíblia | [bíblia] |
| | padre | = PAL | badré | [badré] |
| | farol | = PAL | baróor | [baróor] |
| | sábado | = PAL | sebadóng | [səbaðóŋ] |
| Ger. | Bild | = PAL | bilt | [bilt$^{h}$] |
| | Post | = PAL | bost | [bost$^{h}$] |
| | Papier | = PAL | babiér | [babiɛ́r] |
| Eng. | piston | = PAL | bistóng | [bistóŋ] |
| | box | = PAL | baks | [baks] |
| | valve | = PAL | barb | [barp] |
| | table | = PAL | tébel | [tébəl] |

(6) [p]

(i) adjacent to consonant other than l in same syllable:

Ger. Schlips = PAL slibs [slips]

Eng. present = PAL breséngt [prɛsɛ́nt$^{h}$]

softball = PAL sabtbol [saptból]

(ii) word-finally:

| | | | | |
|---|---|---|---|---|
| Ger. | Brief | = PAL | berí**b** | [bəríp] |
| | Schraube | = PAL | seráu**b** | [səráwp] |
| Eng. | pipe | = PAL | bá**e**b | [báɛp] |
| | club | = PAL | kla**b** | [klap] |

A few exceptions exist, of course, to the patterns of loanword phonology described above. In several borrowings from Eng., the labio-dental fricative **f** is not changed into **/b/**, but instead remains as [f]. PAL words showing this innovating pronunciation are, for example, **féngda** *fender*, and **taifún**[9] *typhoon*. In addition, the initial and medial p's of Sp. Papa *pope* are also pronounced [p] in PAL **Papa**.[10] The PAL word **bostól** *apostle* is a rendition of Sp. apóstol, but the initial **a-** was evidently misinterpreted as the PAL prenominal particle **a**, resulting in **bostól** being taken as the independent noun.[11] The PAL noun **Sebangiól** *Spain* is derived from the Sp. adjective español *Spanish*; in this word, pronounced [**səbaŋyól**], the initial vowel e [ɛ] has been dropped and a [ə] has been inserted to avoid a rather infrequent PAL word-initial consonant cluster **sb-**. Finally, in a word like PAL **nebtái** from Eng. *necktie*, we observe an unusual switching from one grave consonant to another (k to b).

The Jp. bilabial stops p[12] and b are borrowed into PAL in the expected way — that is, they are rendered as the appropriately conditioned allophone of PAL **/b/**. Thus, the examples below require no special explanation:

(7)

| | Loan Words | | PAL | |
|---|---|---|---|---|
| Jp. | bunpō | *grammar* | búmbo | |
| | denpō | *telegram* | démbo | |
| | hinpyōkai | *exhibition, fair* | himbiókai | |
| | kenpei | *military policeman* | kembéi | *police* |
| | senpūki | *electric fan* | sembukí | |
| | nappa | [*dialect*] *greens* | náppa | *(long) cabbage* |
| | bara | *rose* | bára | |
| | bengoshi | *lawyer* | bengngós | |
| | bun | *minute, part* | bung | |
| | ashiba | *scaffolding* | chásiba | |
| | dobu | *ditch* | dóbu | |
| | kyabu[13] | *carburetor* | kiáb | |

In Jp. the phoneme /h/ has three allophones, distributed as follows: before **u** in the syllable /hu/ (written as fu in the Romanisation used here), /h/ is pronounced as a voiceless bilabial fricative [Φ]; before i in the syllable /hi/ and before y in /hya, hyu, hyo/, it is pronounced as a voiceless palatal or prevelar fricative [ç] similar to the fricative in standard Ger. ich *I*);[14] and before other vowels in the syllables /ha, he, ho/, it is a glottal fricative similar to Eng. /h/. Although there is one rare example — PAL kangngob (< Jp. **kangofu** *nurse*) — in which an allophone of Jp. /h/ has become an allophone of PAL /b/, the allophones of Jp. /h/ are normally borrowed into PAL as a new phoneme /h/ with a single allophone, the (voiceless) glottal fricative [h].[15] The following PAL loanwords containing /h/ ([h]) are subclassified according to the original Jp. syllable containing the /h/:

| (8) | | Loan Source | | PAL | |
|---|---|---|---|---|---|
| ha [ha] | haizara | *ashtray* | haisára | |
| | hanafuda | *Jp. card game* | hanahúda | |
| | hashi | *chopsticks* | hási | |
| | hantai | *opposite* | hantái | |
| | harau | *pay (for)* | haráu | |
| hi [ç̣i] | hinpyōkai | *exhibition, fair* | himbiókai | |
| | himitsu | *secret* | hímits | |
| | kōhī | *coffee* | kohí | |
| fu [Φu] | ifukuro | *paunch of stomach* | chihúkuro | |
| | furansu | *France* | huráns | |
| | fūsen | *balloon* | huséng | |
| | fūtō | *envelope* | húto | |
| | futsū | *common, usual* | hútsu | |
| | mafurā | *muffler, scarf* | mahurá | |
| he [he] | henji | *answer, reply* | henzí | |
| ho [ho] | hokori | *dust* | hokorí | |
| | hon | *book* | hong | |
| | hontō | *main island* | hónto | *Babeldaob* |
| | mahōbin | *thermos bottle* | mahóbing | |
| | mihon | *sample* | míhong | |

Until the large influx of Jp. vocabulary, /h/ was not really well established as a phoneme of PAL. During the Sp. period, a few words had been borrowed in which Sp. (orthographic) g or j, both pronounced as the voiceless velar fricative [x], were rendered in PAL as [x] or [h]. These two phones are in free variation even today in such words as the following:[16]

| (9) | | Loan Source | PAL | |
|---|---|---|---|---|
| Sp. | Virgen | *Virgin* | birhén | |
| | angél | *angel* | changhél | |
| | Jesús | *Jesus* | Hesús | |
| | cajón | *box, chest* | káhol | *wooden box, coffin* |

Once /h/ was established as a PAL phoneme after the introduction of many Jp. lexical items, the way was paved for assimilating certain Eng. loanwords containing this sound. Thus, in the words below, Eng. /h/ becomes PAL /h/:

| (10) | Eng. | helmet | hermét |
|---|---|---|---|
| | | home run | homráng |
| | | hose | hos |
| | | hotel | hotér |

In a few cases, however, the Eng. /h/ is lost, as in

| (11) | Eng. | hinge | íings |
|---|---|---|---|
| | | hospital | osbitár |
| | | horse | uós |

2.1.2.

PAL has two contrasting *dental* stops, voiceless /t/ and voiced /d/. The phoneme /t/ has two allophones: word-finally, an aspirated variant [tʰ] occurs, while in all other environments (i.e. word-initially and word-internally) the unaspirated variant [t] is found (compare chat [ʔatʰ] *smoke* with tang [taŋ] *one*). The allophonic variation of /d/ is somewhat more complex, however. Word-initially before a vowel, /d/ is pronounced as either of the voiced allophones [d] or [ð], with the former more likely to occur in rapid, casual speech. Thus, two possible pronunciations are heard for words such as dub [dup, ðup] *dynamite* and deb [dɛp, ðɛp] *sugar cane*. Word-initially before a consonant, the pronunciation of /d/ varies between the voiceless allophones [t] and [θ], with [t] appearing before /b/ and /m/ (e.g. dbak [tbakʰ] *my dynamite*) and [θ] occurring before /k/ and /ng/ (e.g. dngod [θŋoð] *tattoo needle*). Intervocalically and word-finally, /d/ is normally realised as the voiced allophone [ð] (e.g. medak [məðákʰ] *my eye*, bad [bað] *stone*). In final position, /d/ sometimes begins as [ð] and trails off into the voiceless variant [θ].

Dental or alveolar stops from Sp., Ger., Jp., and Eng. are simply borrowed into PAL as the appropriately conditioned variants of /t/ or /d/, and in a few loanwords from Jp. geminate t is also permitted. Note the examples below, which all conform to this principle:

(12)

| | Loan Source | | PAL | |
|---|---|---|---|---|
| Sp. | adiós[17] | *good-bye* | adiós | [aðiós] |
| | diablo | *devil* | diablóng | [diabló̱ŋ, ðiabló̱ŋ] |
| | sandía | *type of watermelon* | sangdiáng | [sanðiáŋ] |
| | soldado | *soldier* | soldáu[18] | [soldáw] |
| Ger. | Turm | *tower* | turm | [turm] *steeple* |
| | Post | *post office* | bost | [bostʰ] |
| | Rad [rat][19] | *bicycle* | rrat | [r̃atʰ] |
| | Schwester | *nun, sister* | suestér | [swɛstér] |
| Jp. | tane | *seed* | táne | [tánɛ] |
| | ten | *grade, point* | teng | [tɛŋ] |
| | tosen | *ferry boat* | toséng | [tosé̱ŋ] |
| | bentō | *box lunch* | bénto | [bénto] |
| | amatte iru | *(be) left over, plenty* | chamattér | [ʔamat:ér] |
| | dai | *platform* | dái | [dái, ðái] |
| | doisu | *Germany* | Dóis | [dóis, ðóis] |
| | bakudan | *bomb* | bakudáng | [baguðáŋ] |
| Eng. | bucket | | báket | [bágɛtʰ] |
| | court | | kort | [kortʰ] |
| | store | | stoáng | [stoáŋ] |
| | time | | táem | [táɛm] |
| | dance | | dangs | [dans, ðans] |
| | bid | | bid | [bið] *auction, bidding* |
| | (re)bound | | bangd | [banð] |

A small, but significant, number of exceptions appear among Eng. loanwords, most of which involve reinterpretation of Eng. word-final (or, less frequently, word-initial) /d/ [d] as PAL /t/. Thus, we find PAL bet [bɛtʰ] (< Eng. bed), kat [katʰ] (< Eng. (playing) card), kolt [koltʰ] (< Eng. gold), rot [rotʰ] (< Eng. rod), and toktáng [toktáŋ] (< Eng. doctor). Two other unusual examples

include PAL bos (< Eng. boat), in which a final stop has been rendered as a fricative, and ballás (< Eng. ballast), in which the word-final consonant cluster has been simplified. Interestingly enough, there are no attested examples of Eng. loanwords containing orthographic th ([θ] or [ð]), so it is not known how these phones would be interpreted in PAL.

## 2.1.3.

PAL has only one indigenous fricative phoneme, the *alveolar* spirant /s/, with a single allophone [s]. There is no voiced counterpart /z/, and no *palatal* fricatives such as /š/ or /ž/. The total absence of palatal articulations in PAL results in several interesting distortions when PAL adopts Sp., Ger., Jp., or Eng. words containing the palatal fricative [š] or the palatal affricates [č] and [ǰ].

To take the simpler cases first, loanwords containing [s] pose no problem when borrowed, as the [s] merely becomes PAL /s/ (with geminate /s/ also being accepted from Jp.). Several representative examples are given below:

| (13) | Loan Source | | PAL | |
|---|---|---|---|---|
| Sp. | Dios | *God* | diós | |
| | cruz[20] | *cross* | kerús | |
| | gracia[21] | *grace* | krásia | |
| | santo | *saint* | sángto | |
| Ger. | Fenster | *window* | béngster | |
| Jp. | sao | *pole, rod* | saó | *fishing pole* |
| | sensei | *teacher* | sénsei | |
| | sumi | *charcoal* | súmi | |
| | bussōge | *hibiscus* | bussónge | |
| Eng. | school | | skúul | |
| | sauce | | sos | |
| | police | | bulís | |

When Jp. words containing the syllables za, zu, ze, and zo are borrowed into PAL, the syllable-initial z — pronounced in Jp. either as the voiced fricative [z] or the voiced affricate [ʒ] (= [d͡ʒ]) — is normally devoiced to PAL /s/ [s]. This change, which is resisted by many speakers of the older generation with a knowledge of Jp., is illustrated in the examples below:

| (14) | Loan Source | | PAL |
|---|---|---|---|
| Jp. | zuga | *drawing, picture* | sungá |
| | zubon | *pants, trousers* | subóng |
| | zurui | *sly, foxy* | surúi |
| | zōri | *rubber thongs* | sóri |
| | zeitaku | *luxurious* | séitak |
| | haizara | *ashtray* | haisára |
| | kanzume | *canned goods* | kansumé |
| | kizu | *injury, scar* | kísu |

A small number of Eng. borrowings with /z/ [z] also exhibit devoicing to PAL /s/ [s] — e.g. PAL bresé́ngt (< Eng. present), PAL hos (< Eng. hose), and PAL kíis *key* (< Eng. keys).[22]

Because PAL lacks any palatal fricatives such as /š/ or /ž/, the nearest equivalent — namely, /s/ [s] — is substituted for the [š] of contributing languages. In Ger. and Eng. /š/ is a separate phoneme whose major allophone [š] is pronounced with concomitant lip rounding. Note the examples below:

| | | | | |
|---|---|---|---|---|
| (15) | Ger. | Schrank | *cupboard, shelf* | serángk[23] |
| | | Schraube | *screw* | seráub |
| | | Schlips | *necktie* | slibs |
| | | Maschine | *machine* | mesí́l |
| | Eng. | shovel | | sébel |
| | | sheep | | sib |
| | | bishop | | bisób |
| | | Marshall (Islands) | | Marsiál |

In Jp., [š] is not an independent phoneme but actually an allophone of /s/ that occurs before the high front vowel /i/ and the palatal glide /y/. This sound, which is not accompanied by lip rounding and is somewhat tenser, higher, and further back than its Eng. counterpart, is found in the Jp. syllables shi, sha, shu, and sho (phonemically /si/, /sya/, /syu/, and /syo/, respectively). It is adopted into PAL as /s/ [s], as the following examples illustrate:

| | | Loan Source | | PAL |
|---|---|---|---|---|
| (16) | Jp. | shashin | *photo* | siasí́ng |
| | | shina | *China* | Sí́na |
| | | shōbai | *business* | sióbai |
| | | shūkan | *custom, habit* | siúkang |
| | | shiken | *test, exam* | skeng[24] |
| | | basho | *place, locality* | básio |
| | | haisha | *dentist* | háisia |
| | | mushiba | *cavity* | musibá |

PAL has adopted a considerable number of loanwords from Jp. containing the affricates [č] and [ǰ]. These affricates (voiceless and voiced, respectively) are articulated in Jp. as dental stops released into palatal fricatives (therefore, they might be more accurately characterised by the phonetic symbolisations [t͡š] and [d͡ž]). These sounds occur in Jp. in the same environments as [š] — namely, before the high front vowel /i/ and the palatal glide /y/ — and represent allophones, respectively, of /t/ and /z/. These allophones are found in the syllables chi, cha, chu, cho (phonemically /ti/, /tya/, /tyu/, /tyo/) and in the syllables ji, ja, ju, jo (usually analysed phonemically as /zi/, /zya/, /zyu/, /zyo/). The Jp. affricates [č] and [ǰ] have had an overt influence on the PAL phonological system, since they have been reinterpreted as the PAL affricates [c] and [ɟ], which are not native to the language. It is only in loanwords from Jp. (and Eng.) that we hear PAL [c] and [ɟ], which are affricates consisting of a dental stop released into an alveolar fricative, perhaps more accurately represented as [t͡s] and [d͡z]. When we compare the original Jp. [t͡š] and [d͡ž] with their PAL counterparts [t͡s] and [d͡z], we notice two important facts. First, the fricative portion of the affricate (a palatal articulation in Jp.) becomes an alveolar fricative in PAL, a natural distortion considering the fact that PAL has no palatal fricatives but only the alveolar fricative /s/ [s]. Second, the development of PAL [d͡z], while filling a gap in the pattern, introduces

the voiced affricate articulation [ z ] into the PAL phonological system as part of the affricate [ d͡z ]. In the examples of (17) below, we find cases of Jp. [ t͡ʃ ] (Romanised ch) = PAL [ t͡s ] (orthographic ts), and in the examples of (18), there are instances of Jp. [ d͡ʒ ] (Romanised j) = PAL [ d͡z ] (orthographic z):

| (17) | Loan Source | | PAL | |
|---|---|---|---|---|
| Jp. | chōshi | *condition, state* | tsiós | |
| | chūbu | *(inner) tube* | tsiúb(u) | |
| | chūi (suru) | *be careful* | tsiuí | |
| | bakuchi | *gambling* | bakutsí | |
| | ochiru | *fall, fail (exam)* | otsír | |
| | denchi | *battery* | dénts(i) | |
| | kimochi | *feeling* | kimóts | |
| | machi | *town, city* | máts(i) | *capital, main town* |
| | kechi | *stingy* | kets | |
| | mōchō | *appendix* | mótsio | *appendicitis* |
| (18) | jakki | *(car) jack* | ziákki | |
| | jū | *gun, rifle* | ziú | |
| | neji | *screw* | nézi | |
| | aji | *taste, flavour* | cházi | |
| | benjo | *toilet* | bénzio | |
| | daijōbu | *all right, OK* | daiziób | |

In (17) above, we have seen examples in which Jp. ch ([ t͡ʃ ]) is rendered as PAL ts ([ t͡s ]). In addition, the Jp. affricate ts [ t͡s ] (an allophone of the phoneme /t/ that occurs only before the high back vowel u), also gets adopted into PAL as ts. This straightforward correspondence is illustrated in the examples below:

| (19) | Loan Source | | PAL | |
|---|---|---|---|---|
| Jp. | tsubame | *barn swallow* | tsubáme | |
| | tsunami | *tidal wave* | tsunámi | |
| | himitsu | *secret* | hímits | |
| | katsuo | *bonito* | kátsuo | |
| | nimotsu | *baggage* | nímots | |
| | dokuritsu (suru) | *become independent* | dokuríts | *independent* |
| | utsusu | *take (photo of)* | chuts(i)ús[25] | |

Looking at loanwords from Eng., we observe that Eng. ch [ t͡ʃ ] as well as the consonant cluster ts is borrowed into PAL as ts [ t͡s ]. Eng. ts corresponds to PAL ts in several cases where the plural of an Eng. noun ending in t has been reinterpreted as a PAL singular (cf. fns. 7 and 22) — e.g. PAL cháuts (< Eng. outs (in baseball)) and PAL dónats (< Eng. doughnuts). Eng. ch becomes PAL ts in items such as tsésa (< Eng. (beer) chaser), tsiokkolét (< Eng. chocolate), and tsuingám (< Eng. chewing gum).[26]

In this complex area of PAL loanword phonology, there are bound to be some exceptions or irregularities. The small, but significant, number of these all involve cases in which [ t͡ʃ ] or [ d͡ʒ ] of the contributing language is rendered as the PAL fricative /s/ [ s ] rather than the expected innovating affricate ts [ t͡s ] or z [ d͡z ]. This simplification is observed in words such as PAL kusaráng (< Sp. cuchara *spoon*), PAL suklatéi *cocoa tree* (< Sp. chocolate *chocolate*), PAL sidósia (< Jp. jidōsha *auto*), PAL sikáng (< Jp. jikan *hour*), PAL íings (< Eng. inch or hinge),[27] and PAL masés (singular) (< Eng. matches).

2.1.4.

PAL has two *velar* consonant phonemes, the stop /k/ and the nasal /ŋ/. The phoneme /k/ has three major allophones, distributed as follows: in word-final position, the aspirated allophone [$k^h$] occurs[28] (e.g. brak [prak$^h$] *taro*, derúmk [ðərúmk$^h$] *thunder*); intervocalically, /k/ is realised as the voiced stop [g] (e.g. rekás [rəgás] *mosquito*, mekeáld [məgɛáld] *hot*); and elsewhere (i.e. word-initially or word-medially adjacent to another consonant) /k/ has the unaspirated allophone [k] (e.g. kar [kar] *medicine*, skors [skors] *cane*. When [k] and [g] are contained in loanwords from Sp., Ger., Jp., and Eng., these sounds are simply converted into the appropriately conditioned allophones of PAL /k/. Thus, the examples below require no further explanation:

(20)

| Loan Source | | | PAL |
|---|---|---|---|
| Sp. | carro | *pushcart, waggon* | karróng |
| | católico | *Catholic* | katolík |
| | Pascua | *Easter* | baskuá |
| | manteca | *lard* | mangtekáng |
| | gracia | *grace* | krásia |
| | iglesia | *church* | ikelésia |
| Ger. | Kapitän | *captain* | kabitéi |
| | Mark | *mark (monetary unit)* | mak |
| | Schrank | *cupboard, shelf* | serángk |
| | Gummi | *rubber* | kúmi |
| Jp. | kai | *shell* | kái |
| | kakine | *fence* | kákine |
| | garasu | *glass* | karás |
| | ginkō | *bank* | kíngko |
| | gomen | *pardon me* | koméng |
| Eng. | clock | | klok |
| | cake | | kéik |
| | tank | | tangk |
| | guitar | | kitá |
| | glove | | kurób |
| | bucket | | báket |

The velar nasal phoneme /ŋ/ (orthographic ng) has two major allophones — [n] and [ŋ] — whose distribution is by and large determined by a principle of assimilation. Before the dental stops /t/ and /d/, the alveolar fricative /s/, and the liquid /r/, the more forward dental variant [n] is used; this allophone is spelled ng in all native PAL words — e.g. iúngs [yuns] *island*, merángd [məránð] *(type of) coral*. Phonemic /n/ [n] is totally lacking in native PAL vocabulary as the consequence of a very early sound change in which proto-Austronesian /*n/ and /*ñ/ became /l/ in various positions except that before /t/, /d/, etc., thus relegating [n] to the status of an allophone of /ŋ/. The resulting resistance to [n] in prevocalic or word-final position was at least partially viable through the Ger. times, as attested from the following borrowings from Sp. and Ger. in which the [n] of the contributing language was taken into PAL as [l]:

(21)

| | Loan Source | | PAL | |
|---|---|---|---|---|
| Sp. | cajón | *box, chest* | káhol | *wooden box, coffin* |
| | cañón | *cannon* | kalióI[29] | |
| | campana | *bell* | kambaláng[30] | |
| | Marianas | *Marianas (Islands)* | Marialás | |
| Ger. | Maschine | *machine* | mesíl | |
| | Grammophon | *phonograph* | karmoból | |

However, we also find examples of Sp. borrowings containing prevocalic [n] in which the [n] was retained in PAL, thus indicating the beginning of a trend to accept this sound in a wider number of environments. Such examples (in which PAL orthographic n is used rather than ng) include keristiáno (< Sp. cristiano *Christian*), komunión (< Sp. comunión *Holy Communion*), and koróna (< Sp. corona *crown*). With the large influx of Jp. vocabulary, prevocalic [n] gradually became established as a possible PAL articulation, and this situation has by and large continued through the present period of Eng. borrowings.[31] In the list below, representative examples are taken from Jp. and Eng. in which prevocalic [n] is adopted without change into PAL:

(22)

| | Loan Source | | PAL |
|---|---|---|---|
| Jp. | nasu | *eggplant* | nas |
| | nikibi | *pimple, acne* | níkibi |
| | negi | *green onion* | néngi |
| | nori | *glue, paste, starch* | nóri |
| | okane | *money* | okáne |
| | tane | *seed* | táne |
| | abunai | *dangerous* | chabunái |
| Eng. | navy | | néibi |
| | nurse | | nurs |
| | gardenia | | kadénia |
| | tennis | | ténis |

In borrowed words containing [n] before dentals or alveolars, PAL simply uses the [n] allophone of /ŋ/. By convention, this [n] is spelled ng in borrowings from Sp., Ger., and Eng., but spelled only as n in loanwords from Jp. These potentially confusing orthographic conventions are to be kept in mind when examining the lists below:

(23)

| | Loan Source | | PAL | |
|---|---|---|---|---|
| Sp. | bandera | *flag, banner* | bangderáng | |
| | manta | *black cloth* | mangtáng | *woollen blanket, cotton* |
| | santo | *saint (m.)* | sángto | |
| Ger. | Fenster | *window* | béngster | |
| | auswendig | *by heart* | chausbéngdik | *memorise* |
| Eng. | pound | | bongd | |
| | dance | | dangs | |
| | inch, hinge | | íings | |
| | ground | | kurángd | *playground* |
| | Sunday | | sángdei | *Sunday, week* |

(24)

| | | | |
|---|---|---|---|
| Jp. | bentō | *box lunch* | bénto |
| | denchi | *battery* | dénts(i) |

| | | |
|---|---|---|
| ensoku | *picnic,outing* | ensók |
| mondai | *problem* | móndai |
| ninjin | *carrot* | nínzin *type of sweet potato* |

The second allophone of /ŋ/ — namely, [ŋ] — is used in native PAL words in all positions where the allophone [n] is excluded. In other words, [ŋ] is of much wider distribution than [n] because it occurs prevocalically (both in word-initial and medial positions), word-finally, and before the velar stop /k/, the glottal stop /ʔ/, and the liquid /l/. Some examples of the [ŋ] allophone of /ŋ/ in native words are given below:

| | | | |
|---|---|---|---|
| (25) | prevocalically: | ngau | *fire* |
| | | ngor | *mouth* |
| | | mengílt | *put oil on* |
| | | ongéu | *steering wheel, rudder* |
| | word-finally: | bung | *flower* |
| | | reng | *heart, spirit* |
| | | ding | *ear* |
| | before /k/, /ʔ/, /l/: | ngklek | *my name* |
| | | omengkángk | *lay on ground* |
| | | Ngchésar | *village in E. Babeldaob* |
| | | nglóik | *dance* |
| | | nglái | *taken, brought* |

In PAL words of foreign origin, the sound [ŋ] (always spelled ng) appears from several sources. First of all, in Ger., Jp., and Eng., [ŋ] is probably best analysed not as an independent phoneme, but rather as an allophone of /n/ before velar consonants. Since the combination *velar nasal* [ŋ] + *velar stop* is natural in PAL (cf. 25 above), it is borrowed without modification, as the examples below indicate:

| (26) | Loan Source | | PAL | |
|---|---|---|---|---|
| Ger. | schenken | *send* | sengk | *gift on child's first birthday* |
| | Schrank | *cupboard, shelf* | serángk | |
| Jp. | denki | *electricity* | déngki | |
| | hankachi | *handkerchief* | hangkáts | |
| | hōsenka | *garden balsam* | hoséngka | |
| | kankei | *relationship* | kángkei | |
| | kankōdan | *tourist group* | kangkódang | *tourist* |
| | katorisenko | *mosquito coil* | katoriséngko | |
| Eng. | bank | | bangk | |
| | monkey (card game) | | mongkíi | |
| | sunglasses | | sangklás[32] | |
| | soft-drink | | sobdríngk | |
| | tank | | tangk | |
| | trunk | | torángk | |

In a large number of PAL loanword items, [ŋ] occurs word-finally. One of the major sources of this [ŋ] is a word-final [n] in the contributing language. Thus, Jp. word-final n (phonetically a nasalised version of the preceding vowel and, for many speakers, somewhat velarised) is invariably taken into PAL as [ŋ], as the examples below indicate:

(27)

| | Loan Source | | PAL |
|---|---|---|---|
| Jp. | barikan | *hair clippers* | barikáng |
| | byōin | *hospital* | bioíng |
| | botan | *button* | botáng |
| | bun | *minute, part* | bung |
| | daikon | *radish* | dáikong |
| | yōchien | *kindergarten* | iotsiéng |
| | sen | *line, wire* | seng |
| | udon | *noodles* | udóng |

The same phenomenon is observed in Eng. borrowings, where word-final Eng. [n] is pronounced in PAL as [ŋ]:[33]

(28)

| | | | |
|---|---|---|---|
| Eng. | piston | bistóng | |
| | aspirin | chasberíng | |
| | drum (can) | deromukáng | *water drum* |
| | home run | homráng | |
| | curtain | katéng | |
| | sign | saíng | |

A second major source of word-final [ŋ] in loanwords is clearly due to excrescence. In most native PAL words, an excrescent word-final [ŋ] appears after the vowels a, o, and u when the word in question appears in isolation (e.g. when being cited), in sentence-final position, or sentence-internally before a major pause. To take a typical example, the PAL intransitive verb mong *go* has no excrescent [ŋ] within a sentence, as in (29a), but does take this [ŋ] sentence-finally, as in (29b):

(29)

| | | |
|---|---|---|
| a. | Ng mo er a skuul | *He's going to school.* |
| b. | Ng mong. | *He's going.* |

A good number of Sp. words ending in a and o take an excrescent word-final [ŋ] in PAL. When the word in question is used sentence-internally (not before a pause), the [ŋ] is normally lost. Some representative examples include the following:

(30)

| | | | |
|---|---|---|---|
| Sp. | bandera | *flag, banner* | bangderáng |
| | plato | *dish, plate* | belatóng |
| | botella | *bottle* | butiliáng |
| | diablo | *devil* | diablóng |
| | martillo | *hammer* | martilióng |
| | Misa | *Mass* | Misáng |

Interestingly enough, many loanwords from Jp. end in the vowels a, o, and u, but virtually none of them exhibit excrescent [ŋ].[34] Why Jp. borrowings should have been resistant to this phenomenon is quite mysterious, especially because many Eng. borrowings, which were of course chronologically later, were clearly susceptible to it. In the majority of Eng. loanwords below, note the loss of word-final Eng. [r] and its replacement in PAL by the vowel [a] followed by an excrescent [ŋ]:

| (31) | Eng. | verandah | berangdáng |
|---|---|---|---|
| | | beer | biáng |
| | | flour | blauáng |
| | | master | mastáng |
| | | store | stoáng |
| | | doctor | tóktang |
| | | number | lámbang |

A few exceptions to this pattern exist — e.g. PAL botéto (< Eng. potato), PAL kitá (< Eng. guitar), and PAL tsésa (< Eng. (beer) chaser).

In certain dialects of Jp. such as Tokyo standard, the voiced velar stop /g/ is pronounced between voiced sounds as a nsasl ([ŋ]). Since the standard dialect was used by government officials and taught in the Palau schools during the Jp. period, it is not surprising that most Jp. words with intervocalic /g/ [ŋ] were borrowed into PAL with [ŋ]. Some typical examples are listed here:

| (32) | Loan Source | | PAL |
|---|---|---|---|
| Jp. | agaru | *rise, increase* | changár |
| | dōgu | *tool* | dóngu |
| | yanagi | *willow* | ianángi |
| | maegami | *bangs* | maingami |
| | maguro | *tuna* | manguró |
| | saigo | *last time* | sáingo |
| | tamago | *egg* | tamángo |

In the standard dialect of Jp., when a syllable ending in n is followed by a syllable beginning with g, the resulting sound is something like a long velar nasal — e.g. bangō [baã$^{ŋ}$ŋo:] *member*. A similar long velar nasal is pronounced in the corresponding PAL loanword and is represented by the orthographic sequence ngng:

| (33) | Jp. | bangō | *number* | bángngo |
|---|---|---|---|---|
| | | bengoshi | *lawyer* | bengngós |
| | | kangofu | *nurse* | kangngób |
| | | ringo | *apple* | ríngngo |
| | | songai | *(financial) loss* | sóngngai |

## 2.1.5.

The PAL *glottal* stop phoneme /ʔ/ (orthographic ch) has no allophonic variation, and is pronounced [ʔ] in all positions (word-initial, word-internal, and word-final) — e.g. charm *animal*, chisél *news of him*, mechéd *shallow*, dechór *standing*, táoch *mangrove channel*, and búuch *betel nut*. None of the languages contributing loanwords to PAL has phonemic glottal stop, although in Jp. vowel-initial words are normally pronounced with an excrescent prevocalic glottal stop both in isolation and within sentences. The same phenomenon tends to occur in Eng., especially when vowel-initial items are uttered in isolation. Most Jp. and Eng. loanwords in this category have been borrowed into PAL with word-initial glottal stop, as the following sampling of vocabulary indicates:

(34)

| | Loan Source | | PAL | |
|---|---|---|---|---|
| Jp. | aburasashi | *oil can* | chaburasási | |
| | abunai | *dangerous* | chabunái | |
| | eisei | *sanitation* | chéisei | |
| | imi | *meaning* | chímí | |
| | oto | *noise, sound* | chóto | |
| | uri | *melon, cucumber* | chúri | *muskmelon* |
| | usui | *(liquid) weak* | chusúi | |
| Eng. | ambulance | | chambelángs | |
| | air (for tyre, etc.) | | chéa | |
| | ice | | cháis | |
| | okra | | chókura | |

There are, of course, cases in which the PAL form lacks the initial glottal stop — e.g. PAL ensók (< Jp. ensoku *picnic, outing*), PAL okáne (< Jp. okane *money*), PAL usángi (< Jp. usagi *rabbit*), PAL obís (< Eng. office), and PAL osbitár (< Eng. hospital).

The Jp. syllables wa, ya, yu, and yo, which begin with glides, are adopted into PAL as two separate syllables — that is, w and y become full vowels. When PAL word-initial u or i has developed from this source, there is no word-initial glottal stop. Several representative examples are PAL uáta (< Jp. wata *cotton*), PAL uatasibúne (< Jp. watashibune *ferry boat*), PAL iákiu (< Jp. yakyū *baseball*), PAL iásai (< Jp. yasai *vegetables*), PAL iaksók (< Jp. yakusoku *promise*), and PAL iótei (< Jp. yotei *plan, schedule*).

## 2.1.6.

Palauan has two *liquids* with phonemic status — /l/ and /r/. The phoneme /l/ is a lateral articulation similar to Eng. /l/: it shows no allophonic variation and can occur 'long' ([l:] = phonemic /ll/ and orthographic ll). The phoneme /r/ is a flap r [r] very similar to Jp. /r/, and when two r's occur phonemically in sequence (/rr/ = orthographic rr) the result is a trilled r, not unlike Sp. [r̃] but often pronounced with considerable friction. In spite of the fact that PAL has both /l/ and /r/ in contrast (e.g. lisél *its coconut tree* vs. risél *its root*), these phonemes are often interchanged in loanwords, as we will see below. The instability of an absolutely clearcut contrast between /l/ and /r/ is evidenced not only by such loanwords but also by certain native pairs of free variants (e.g. merrédel — merréder *leader* and ĺikr — ĺikl *outside*) exhibiting an interchange of these sounds. While there are a few examples in which r of the contributing language is taken into PAL as l, in the great majority of cases non-indigenous l is adopted as PAL r:[35]

(35)

| | | | | |
|---|---|---|---|---|
| Sp. | faról | *lantern* | barór | |
| | sal | *salt* | sar | |
| | azucár | *sugar* | sukál[36] | |
| Ger. | Tafel | *blackboard* | tabér | |
| | Papier | *paper, document* | babiér — babilngél[37] | *paper, letter, book* |

| | | | |
|---|---|---|---|
| Eng. | valve | barb | |
| | belt | bert | *fan belt* |
| | ball (in baseball) | bor | |
| | helmet | hermét | |
| | hospital | osbitár | |
| | hotel | hotér | |
| | keel | kir[38] | |
| | glove | kurób | |

Interestingly enough, when words are borrowed from Jp., which has only /r/ and no /l/, there is never any distortion: all Jp. /r/ [r] become PAL /r/ — e.g. PAL chíro (< Jp. iro *colour*).

Even though PAL has word-final r and r before other consonants (e.g. ngor *mouth*, skors *cane*, dart *one hundred*), when r occurs in these positions in loanwords, it is often lost. Thus, Ger. Mark *mark (monetary unit)* becomes PAL mak *fifty cents*, and r is lost word-finally in the Eng. words of (31) and before a consonant in the following additional items:

(36)

| | Loan Source | PAL | |
|---|---|---|---|
| Eng. | carcinoma | kasinóma | *cancer* |
| | card | kat | *playing cards* |
| | market | makít | |
| | horse | uós | |

It is not known to what extent, if any, exposure to Eng. r-less dialects might have influenced the pronunciation of the PAL words in (31) and (36).

The phoneme sequences /ll/ and /rr/ in loanwords also reveal some distortions. Thus, Sp. barril *barrel* shows a lengthening of the l in PAL barríll, and while this very same word and karrong (< Sp. carro *pushcart, wagon*) show a correspondence between PAL /rr/ [ř] and Sp. /ř/, in loanwords rrat (< Ger. Rad *bicycle*) and rrom *liquor* (< Eng. rum) we find an idiosyncratic doubling of the r.

## 2.2. PAL vowels and their use in loanwords

The following chart presents the six vowels of PAL classified in terms of tongue height and relative tongue advancement:

Chart of PAL vowels

| Tongue height | Tongue Advancement: front | central | back |
|---|---|---|---|
| high | i | | u |
| mid | ɛ | ə | o |
| low | | a | |

Figure 2

There are five full vowels in PAL — /i/, /u/, /ɛ/ (orthographic e), /o/, and /a/ — which occur in stressed or unstressed syllables. In addition, there is one reduced vowel — [ə] (also orthographic e) — whose occurrence is restricted to unstressed syllables (very much like Eng. [ə] *schwa*). Vowels in loanwords are taken into PAL according to the following general principle: full vowels in the contributing language are normally pronounced as the PAL vowel closest to them in articulation unless they happen to fall in an unstressed syllable of the PAL word, in which case they are sometimes pronounced as [ə].[39]

## 2.2.1.

Both Sp. and Jp. happen to have, like PAL, a simple system of five full vowels — namely, i, u, e, o, and a. Unlike PAL, however, they have no reduced vowel such as [ə] that occurs in unstressed syllables. Because there is no 'conflict' with the PAL five-vowel system, loanwords from Sp., and Jp. are only rarely taken into PAL with any modification of vowel quality. The only cases of distortion involve vowel raising in Sp. loanwords — e.g.

(37)

| Loan Source | | | PAL | |
|---|---|---|---|---|
| Sp. | botella | *bottle* | butiliáng | (o>u, e>i) |
| | cebollas | *onions* | sebuliás | *onion* (o>u) |
| | chocolate | *chocolate* | suklátei | *cocoa tree* (o>u) |

In addition, full vowels in several Sp. words are reduced to [ə] in an unstressed syllable of the PAL form — e.g. PAL kelebús [kələbús] (< Sp. calabozo *jail, prison*) and PAL sebadóng [səbaðóŋ] (< Sp. sábado *Saturday*). By contrast, the full vowels of Jp. loanwords are never reduced to [ə] even when they occur unstressed in the corresponding PAL form, as many previous examples will attest.

Although the vowel system of Ger. is much more complex than that of PAL, there is little indication of distortion during borrowing. The lack of loanwords containing, for example, the front rounded vowels ü and ö of Ger. is probably due to the small statistical sample (since at most 25 loanwords from Ger. are attested). At any rate, the only phenomenon of interest we find in loanwords from Ger. involves the reduction of full vowels to [ə] in an unstressed syllable of the PAL form — e.g. PAL mesíl [məsíl] (< Ger. Maschine *machine, motor*).

The Eng. vocalic system is very rich, with a relatively large number of single vowels (e.g. in the author's dialect, [I, i, U, u, ɛ, e, ɔ, o, ʌ, a, æ]) and several diphthongs ([aw, ay, ɔy]). When the Eng. vocalic system is 'translated' into PAL, numerous changes and distortions occur, and we find quite a significant number of exceptions and irregularities. The major modifications to be discussed below all involve cases in which a vowel of Eng. (e.g. [æ, ʌ, ɔ]) has no direct PAL equivalent and is therefore rendered as a PAL vowel contiguous to it in terms of articulation.

The Eng. low front vowel [æ] is almost always borrowed as PAL low central [a], as the list below illustrates:

(38)

| | | Loan Source | PAL |
|---|---|---|---|
| Eng. [æ] | > | band | bangd |
| PAL [a] | | valve | barb |
| | | dance | dangs |
| | | stamp | stamb |

Only in one unusual case is Eng. [æ] taken into PAL as [ɛ] — i.e. PAL bek (< Eng. bag).

The Eng. mid central (stressed) vowel [ʌ] is borrowed either as PAL low central [a] or PAL mid back [o], in each case preserving one of the positional features of the original Eng. vowel. Examples of both of these situations are given below:

(39)

| | | | Loan Source | PAL |
|---|---|---|---|---|
| Eng. | [ʌ] | > | bucket | báket |
| PAL | [a] | | doughnuts | dónats |
| | | | club | klab |
| Eng. | [ʌ] | > | pump | bomb |
| PAL | [o] | | cup | kob |
| | | | rum | rrom |

The Eng. lower mid back vowel [ɔ] is borrowed as the PAL mid back vowel [o], as shown below:

(40)

| | | | Loan Source | PAL |
|---|---|---|---|---|
| Eng. | [ɔ] | > | ball | bor |
| PAL | [o] | | sauce | sos |

Eng. has several pairs of contrasting lax vs. tense vowels — i.e. [I] vs. [i], [U] vs. [u], and [ɛ] vs. [e].[40] Since PAL has only one vowel corresponding to each of these pairs, the Eng. lax vs. tense contrast is neutralised in loanwords, as the following data illustrate:

(41)

| | | | | | |
|---|---|---|---|---|---|
| Eng. | [I, i] | > | bid | [bId] | bid |
| PAL | [i] (tense) | | piston | [pístən] | bistóng |
| | | | sheep | [šip] | sib |
| | | | keel | [kil] | kil |
| Eng. | [U, u] | > | puss | [pUs] | bus |
| PAL | [u] (tense) | | pool | [pul] | bul |
| | | | room | [rum] | rum |
| Eng. | [ɛ, e] | > | bed | [bɛd] | bet |
| PAL | [ɛ] (lax) | | tennis | [tɛ́nIs] | tenís |
| | | | save | [sev] | seb |
| | | | chaser | [čésər] | tsésa |

In addition to the above, some Eng. tense vowels become long vowels or diphthongs in PAL — e.g. PAL kíis [ki:s] (< Eng. keys), PAL skúul [sku:l] (< Eng. school), PAL téib [tɛyp] (< Eng. tape), and PAL kéik [kɛyk$^h$] (< Eng. cake). Furthermore, some full vowels or diphthongs of Eng. are reduced to PAL [ə] in unstressed syllables — e.g. PAL chambelángs [ʔambəláns] (< Eng. ambulance) and PAL chasberíng [ʔasbəríŋ] (< Eng. aspirin).

The Eng. diphtnongs [aw] and [ay] are adopted into PAL with much irregularity. In some cases, Eng. [aw] remains a diphthong in PAL (e.g. PAL sáusab [sáwsap] (< Eng. soursop)), while in others it becomes the single vowel a (e.g. PAL kurángd (< Eng. ground)) or the single vowel o (e.g. PAL bongd (< Eng. pound)). Eng. [ay] is either pronounced as a similar diphthong in PAL (e.g. PAL cháis [ʔays] (< Eng. ice)) or as the diphthong [aɛ] (e.g. PAL báeb [báɛp] (< Eng. pipe)).

Some of the more unusual PAL renditions of Eng. vowels are undoubtedly due to spelling pronunciations. Thus, in words like the following, the PAL vowel corresponds to the Eng. orthographic vowel rather than to the actual Eng. phonetic vowel:

| (42) | Loan Source | PAL |
|---|---|---|
| Eng. | pocket | bokkét |
| | bishop | bisób |
| | altar | chaltár |
| | deacon | díakon |
| | clock | klok |
| | compass | kómbas |

## 2.2.2.

In a small number of loanwords from Sp., Ger., and Eng., the PAL form shows an excrescent [ə] between the two consonants that form a cluster in the contributing language. Since kr, sr, and dr are not possible as word-initial consonant clusters in PAL, it is natural that a [ə] would be inserted in the PAL form to obviate an unpronounceable sequence. Thus, cr- initials from Sp. are rendered in PAL as kər- (e.g. PAL keristiáno (< Sp. cristiano *Christian*) and PAL kerús (< Sp. cruz *cross*)), and schr- [šr] initials from Ger. become PAL sər- (e.g. PAL serángk (< Ger. Schrank *cupboard, shelf*) and PAL seráub (< Ger. Schraube *screw*)); in addition, Eng. drum (can) becomes PAL deromukáng. An excrescent [ə] may appear as well, however, between the consonants of clusters that are themselves acceptable in native PAL words — e.g. kl, sb, and br. This unpredictable and sporadic phenomenon is observed in such examples as PAL ikelésia (< Sp. iglesia *church*), PAL Sebangiól *Spain* (< Sp. español *Spanish*), and PAL beríb (< Ger. Brief *letter*).

## 2.2.3.

In Jp. the high vowels i and u are characteristically devoiced between voiceless consonants and in word-final position after a voiceless consonant.[41] Since such voiceless vowels have a very low audibility, it is not surprising that they are very often omitted in PAL loanwords. In the words of (43) below, the omitted voiceless vowel of Jp. is interconsonantal, in the examples of (44) it is word-final, and in (45) it occurs in both positions:

| (43) | Loan Source | | PAL | |
|---|---|---|---|---|
| Jp. | bokusō | *grass, pasture* | boksó | *elephant grass* |
| | kuse | *habit* | kse | |
| | rekishi | *history* | réksi | |
| | shitagi | *underwear* | stángi | *petticoat* |
| | shiken | *test, exam* | skeng | |
| | hikōki | *aeroplane* | skóki[42] | |
| | tsukemono | *pickles* | (t)skemóno[43] | |

| | | | |
|---|---|---|---|
| (44) | bengoshi | *lawyer* | bengngós |
| | doisu | *Germany* | Dóis |
| | doku | *poison* | dok |
| | hachi | *bee* | hats |
| | himitsu | *secret* | hímits |
| | kangofu | *nurse* | kangngób |
| | keikaku | *plan* | kéikak |
| | shokuminchi | *colony, settlement* | siokumíns[44] |
| (45) | mokuteki | *purpose* | mokték |
| | shikaku | *square* | skak |
| | yakusoku | *promise* | iaksók |

In most PAL loanwords from Jp., the high back vowel u is deleted after r; u is also deleted sporadically after certain other voiced consonants. Examples of this phenomenon are illustrated below:

| (46) | Loan Source | | PAL | |
|---|---|---|---|---|
| Jp. | taoru | *towel* | táor | |
| | sarumechīru | (Jp. brand name) | sarmetsír | *liniment* |
| | amatte iru | *left over, plenty* | chamattér | |
| | kiite iru | *effective* | kitér | |
| | kotowaru | *refuse* | kotouár | |
| | tebukuro | *glove* | teb(u)kuró | |
| | daijōbu | *all right, OK* | daiziób | |

## 2.2.4.

Of the five full vowels of PAL, four can occur long (orthographically, ee, ii, oo, and uu). Phonetically, these long vowels are not only greater in length than the corresponding single vowels but also are followed by a slight gliding articulation, at least for many speakers. Native words containing long vowels include kméed [kmɛ:ð] *near*, díil [ði:l] *abdomen*, dekóol [ðəgó:l] *cigarette*, and búuch [bu:ʔ] *betel nut*.[45] As we observed in 2.2.1. above, some PAL loanwords from Eng. contain long vowels as renderings of Eng. tense vowels — to repeat one example, PAL kíis [ki:s] (< Eng. keys). Given the fact that PAL has long vowels both in the native phonology and in Eng. loanword phonology, it is very surprising that there has been a total resistance to borrowing long vowels from Jp. Thus, even though Jp. has long equivalents for all five vowels (Romanised as ā, ī, ū, ē, and ō), there is not a single case in which PAL has not shortened a corresponding Jp. long vowel. A few of the numerous examples of this across-the-board, yet puzzling, phenomenon are provided below:

| | | | | |
|---|---|---|---|---|
| (47) | Jp. | sarumechīru | (Jp. brand name) | sarmetsír *liniment* |
| | | kōhī | *coffee* | kohí |
| | | kyūri | *cucumber* | kiúri |
| | | senpūki | *electric fan* | sembukí |
| | | rajiētā | *radiator* | raziéta |
| | | kōsui | *perfume* | kosúi |

## 3. PAL LOANWORD MORPHOLOGY

The PAL morphological system has been surprisingly flexible in the extent to which it has allowed loanwords to be integrated, as illustrated by the fact that many borrowed noun and verb stems participate in characteristic native inflectional and derivational patterns. In addition, lexical items from diverse grammatical classes in the contributing languages have been readily borrowed into PAL as members of the several major native parts of speech, often accompanied by interesting shifts of meaning.

### 3.1. Change of part-of-speech affiliation during borrowing

In the majority of cases, as we would expect, PAL has borrowed a foreign lexical item as a member of the indigenous grammatical class that most closely corresponds to the original class membership in the contributing language. Thus, most borrowings that are nouns in Sp., Ger., Jp., and Eng. also function as nouns in PAL. To take only one typical example, the loanword sensei *teacher*, a noun in Jp., also is used as a noun in PAL and can occur in three basic nominal positions — i.e. as sentence subject, sentence object, and after the preposition er *of, to, etc.* Each of these is illustrated in the sentences below:

(48) a. A sensei a ungil. *The teacher is good.*

b. Ak milsa a sensei. *I saw the teacher.*

c. Tia a hong er a sensei. *This is the teacher's book (= book of the teacher).*

#### 3.1.1.

Whereas most non-indigenous nouns are borrowed as PAL nouns, a significant number are also adopted into PAL as ***state verbs***. As opposed to ***action verbs***, which designate actions, activities, or events that involve an active participant, PAL state verbs describe states, qualities, or conditions that temporarily or permanently characterise persons or things. Syntactically, action verbs can be distinguished from state verbs in terms of the pattern of past tense formation: action verbs use an infix -il- (e.g. mengaus *weave* — milengaus *was weaving*), while state verbs use the auxiliary mle *was, were* (e.g. ungil *good* — mle ungil *was good*). PAL state verbs normally correspond to Eng. adjectives (and, as we will see below, to various classes of modifying words in Jp.), but not always: thus, for instance, the PAL state verb medengei *know* (cf. past mle medengei *knew*) has an Eng. verb (know) as its nearest equivalent.[46]

Most of the cases in which a foreign noun is borrowed into PAL as a state verb come from Jp. Because the part-of-speech affiliation is changed, the meaning of the lexical item is also modified, but in a fairly consistent manner: thus, as the glosses for the PAL state verbs in (49) below indicate, the PAL meaning is something like 'characterised by/similar in appearance to [the referent of the original Jp. noun]'.

(49)

| | Jp. Noun | PAL State Verb | |
|---|---|---|---|
| baikin | *bacillus, bacterium, germ* | baikíng | *unsanitary, unhygienic* |
| bozu | *Buddhist priest, monk, shaven head* | bózu | *(head) completely shaved* |
| asebo[47] | *prickly heat, heat rash* | chásebo | *broken out in prickly heat* |
| aji | *taste, flavour* | cházi | *tasty* |
| haibyō | *lung disease, TB* | háibio | *sick with TB* |
| hanbun | *half* | hámbung | *half-witted, simple-minded* |
| kama | *sickle, hook* | kámang | *(arm) twisted/crippled* |
| jiman | *pride, vanity* | simáng | *vain, boastful* |
| tamanegi | *onion* | tamanéngi | *(head) completely shaved* |
| tamagogata | *ovoid figure* | tamangongáta | *egg-shaped, oval* |

In addition to being used in PAL as state verbs, some of the items in (49) can also be used as nouns (e.g. cházi *taste, flavour*, kámang *sickle*, and tamanéngi *onion*).

A few examples following the pattern of (49) are found from Sp. and Eng. sources. These are given below:

(50)

| | | Noun | PAL State Verb | |
|---|---|---|---|---|
| Sp. | diablo | *devil* | diablóng | *terrible, awful* |
| | calabozo | *dungeon, cell* | kelebús | *jailed, in prison* |
| | martillo | *hammer* | martilióng | *clumsy, ungraceful* |
| Eng. | curve | | kab | *curved* |
| | butterfly | | báterflai | *(person) fickle/prone to changing one's mind* |
| | you-drive (car) | | iudoráibu | *(woman) loose/fast.*[48] |

All of the words in (50) except báterflai can also be used in PAL as nouns.

Several PAL loanwords illustrate other types of shifts in part-of-speech affiliation resulting from the process of borrowing. Thus, the two Eng. adjectives *English* and *American* are the source for the PAL nouns Ingklís *England* and Merikél *America*.[49] In addition, a few Jp. nouns were adopted into PAL as intransitive verbs — e.g. PAL v.i. chanzáng *add, do sums* (< Jp. n. anzan *mental arithmetic/calculation*) and PAL v.i. sbíido *speed up* (< Jp. n. supiido *speed*). As a final example of interest, the Ger. adverbial auswendig *by heart/memory* is turned into a transitive verb in PAL — i.e. chausbéngdik *know/learn thoroughly, memorise*.

## 3.1.2.

There are three major classes of modifying words in Jp., all of which are borrowed into PAL as state verbs. Jp. *adjectivals* are forms that inflect for tense and other categories: for example, in the present tense, they take the suffix -i and in the past tense the suffix -katta (e.g. usui *(liquid) weak* — usukatta *(liquid) was weak*). When Jp. adjectivals are adopted into PAL they are invariably taken in the present tense (-i) form, as the examples below illustrate:

| (51) | Jp. Adjectival | | PAL State Verb |
|---|---|---|---|
| | abunai | *dangerous* | chabunái |
| | usui | *(liquid) weak* | chusúi |
| | komakai | *detailed, thorough, stingy* | komakái |
| | koi | *(liquid) strong* | kói |
| | kusai | *bad-smelling* | ksái |
| | omoshiroi | *interesting, funny* | omosiróí[50] |
| | sabishii | *lonely* | sabisí |
| | zurui | *sly, foxy, sneaky* | surúi |
| | takai | *expensive* | takái |

Jp. ***nominal adjectivals*** are non-inflecting noun-like elements that are divided into two subclasses according to whether they must be linked to a following modified noun by the particle na or the particle no (e.g. kantan na koto *simple matter* vs. futsū no koto *usual matter*). Nominal adjectivals from both subclasses are taken into PAL as state verbs, as shown in the examples of (52) below:

| (52) | | Jp. Nominal Adjectival | | PAL State Verb | |
|---|---|---|---|---|---|
| | na-type | bonkura | *dull, slow-witted* | bongkurá | |
| | | iroiro | *various, diverse, miscellaneous* | chirochíro | *many-coloured, (children of particular woman) fathered by different men* |
| | | otenba | *pert, saucy* | chotémba | *(woman) loose/fast/ flirtatious* |
| | | daijōbu | *all right, OK* | daiziób | |
| | | kantan | *simple, brief* | kantáng | |
| | | kechi | *stingy, miserly* | kets | |
| | | zeitaku | *luxurious, high-class* | séitak | |
| | no-type | osoroi | *(clothes, etc.) of uniform style* | chosorói | |
| | | daitai | *general, main, rough* | daitái | *all right, OK*[51] |
| | | hadaka | *naked* | hadaká | *naked, bare-breasted* |
| | | hadashi | *barefooted* | hadási | |
| | | hantai | *opposite, dissenting* | hantái | |
| | | futsū | *common, usual* | hútsu | |
| | | katate | *(with) one hand* | kataté | *(with) one hand, dexterous* |
| | | kichigai | *crazy, insane* | kitsingái | |

Many stative expressions in Jp. consist of a verb in its gerund form (with the suffix -te) followed by the auxiliary verb iru *be, exist* — e.g. tsukarete iru *(be) tired* from tsukareru *get tired*). A few of these -te iru expressions have been borrowed into PAL as state verbs: in the resulting PAL forms, of course, the identity of the separate Jp. morphemes (verb stem, suffix -te, and auxiliary verb iru) has been obscured. Some representative examples of this interesting phenomenon are listed below:

(53)

| Jp. Stative Expression | | PAL State Verb |
|---|---|---|
| amatte iru | *(be) left over/plenty, more than enough* | chamátter |
| atte iru | *(be) suitable/appropriate (for each other)* | chattér[52] |
| kankei shite iru | *(be) related to/ connected with* | kangkeistér |
| kiite iru | *(be) effective/ efficacious* | kitér |
| tsukarete iru | *(be) tired* | (t)skáreter |

A small number of Jp. intransitive verbs have been converted into PAL state verbs. In every case, the Jp. intransitive verb is of a special type consisting of a noun followed by the verb suru *do*. In the PAL form, however, the verb suru is omitted:

(54)

| Jp. Intransitive Verb | | PAL State Verb | |
|---|---|---|---|
| dokuritsu (-suru) | *become independent* | dokuríts | *independent* |
| koshō (-suru) | *break down, go out of order* | kósio | *broken, out of order* |
| panku (-suru) | *get punctured, blow out* | bangk | *(tyre) flat/punctured*[53] |

### 3.1.3.

Most transitive and intransitive verbs from Japanese and English are borrowed into PAL without any change in part-of-speech affiliation. Thus, the examples in (55-6) below are rather straightforward, with the semantic changes being of most interest:

(55)

| | Jp. or Eng. Transitive Verb | | PAL Transitive Verb | |
|---|---|---|---|---|
| Jp. | harau | *pay for* | haráu | |
| | yakusoku (-suru) | *promise* | iaksók[54] | |
| | kaburu | *put on/wear (hat)* | kabúr | *flip (person) over one's shoulder and throw down* |
| | kensa (-suru) | *inspect, examine, (medically)* | kénsa[54] | |
| | kotowaru | *refuse* | kotouár | |
| | mawasu | *turn, screw* | mauás | |
| | shimeru | *close, strangle, choke* | simér | *strangle, choke, turn off (e.g. water)* |
| | shiraberu | *examine, investigate* | sirabér | |
| | chūi (-suru) | *be careful, watch out for* | tsiuí | *watch out for (one's behaviour), warn (person)* |
| | ōen (-suru) | *support, cheer* | o(i)éng | *praise, honour, acclaim* |
| | tsukamaeru | *catch, seize, arrest* | (t)skamáer | *confront, face, corner* |

| | | |
|---|---|---|
| Eng. | stop | stob |
| | sign | saíng[55] |
| | bake | béik |

(56)

| | Jp. or Eng. Intransitive Verb | | PAL Intransitive Verb | |
|---|---|---|---|---|
| Jp. | bakuhatsu (-suru) | *blow up, explode* | bakuháts | |
| | ayamaru | *apologise to* | chaiamár | |
| | katsu | *win* | kats[56] | |
| | kawaru | *change* | kauár | *(condition of wind) change* |
| | kizetsu (-suru) | *faint* | kiséts | |
| | makeru | *lose* | máke[57] | |
| | mawaru[58] | *turn/go around* | mauár | |
| | mōkaru | *make profit, gain* | mokár | |
| | shinpai (-suru) | *worry* | símbai | |
| | son (-suru) | *lose money, incur loss* | song | |
| | seikō (-suru) | *succeed* | séiko | |
| | taoreru | *fall, collapse* | taorér | |
| Eng. | drive | | doráib | *drive around in car* |
| | go ahead | | kohéi | *go ahead/forward, advance* |

As exceptions to the above, a few intransitive verbs or expressions of Jp. have been reinterpreted as PAL transitive verbs. The most striking examples are PAL v.t. changár *promote* (< Jp. v.i. agaru *rise, increase*)[59] and PAL v.t. kingátsku *notice, be aware of* (< Jp. v.i. ki ga tsuku *take notice of, be attuned to*).[60]

## 3.2. Paradigmatic features of loanwords

PAL native morphological structure is extremely varied and rich, with highly developed systems of inflection and derivation. Transitive verbs occur in both imperfective and perfective forms, with each type composed of characteristic morphemes and the perfective forms involving a whole paradigm of incorporated pronominal object suffixes. In addition, transitive verbs have related stative forms (the so-called resulting and anticipating state verbs to be explained in 3.2.1. below) as well as derived reciprocals, causatives, and numerous reduplicated forms. Intransitive verbs and state verbs have inchoative forms ('is beginning to..., is becoming...') and predictive forms ('is about to (become)...') and also exhibit derived reciprocals, causatives, and reduplicatives. Finally, the great majority of nouns are inflected with suffixes that indicate the person and number of the possessor (see 3.2.2. below).

### 3.2.1.

While complete coverage of the derivational morphology of PAL verbs is beyond the scope of this paper,[61] the following illustration will suffice to describe those categories relevant to loanword morphology. Given a basic noun

stem such as chelébed *whip, club, anything to hit with,* the following forms can be derived:

| | | | | | | |
|---|---|---|---|---|---|---|
| (57) | a. | *ergative* | mechelébed | *gets hit* | milechelébed | *got hit* |
| | b. | *imperfective* | mengelébed | *hits, is hitting* | milengelébed | *was hitting* |
| | c. | *perfective* | cholebedák | *hits me* | chillebedák | *hit me (past)* |
| | | | cholebedáu | *hits you (sg.)* | chillebedáu | *hit you (sg.)* |
| | | | cholebedíi | *hits him/her/it* | chillebedíi | *hit him/her/it* |
| | d. | *resulting state verb* | chellébed | *(in a state of having been) hit* | | |
| | e. | *anticipating state verb* | chelebedáll | *is/needs to be hit* | | |

Both the ergative and imperfective forms begin with the verb marker me- (o- in verbs formed from b-initial stems) and inflect for the past tense by infixing the past tense marker -il- after the word-initial consonant. In the imperfective form, an additional morpheme — the imperfective marker ng (l or m in other environments) — appears before the stem chelébed, whose initial consonant has been deleted.[62] The perfective forms consist basically of the stem chelébed (but note the change of vowel: the o of cholebed- is most likely a phonetically modified, infixed form of the verb marker) followed by a series of suffixed pronominal objects — e.g. -ák *me,* -áu *you (sg.),* -íi *him/her/it,* and so on. The past tense of perfective forms also involves infixing the past tense marker -il- after the word-initial consonant. The two state verbs — resulting and anticipating — are used as nominal modifiers with rather specialised meanings. The resulting state verb, formed by infixing -(e)l- after the word-initial consonant, describes the state or condition that some person or object is in as a result of having undergone the related action, while the anticipating state verb, formed by suffixing -áll (or -(e)l), focuses on some person or object as being expected or required to undergo that action.

A significant number of loanwords have been Palauanised to the extent that they participate in paradigms similar to (57). Some representative examples are given below in the same format:

(58) PAL n. kíis *key* (< Eng. keys)

a. *erg.* mekíis *gets opened* milekíis *got opened*

b. *impf.* mengíis *opens (with key), unlocks* milengíis *was opening*

c. *pf.* kiisíi *opens it up* kilisíi *opened it up*

d. *res. state vb.* klíis *opened*

e. *antic. state vb.* kisáll *is to be opened*

(59) PAL n. chasuárt *asphalt* (< Eng. asphalt)

a. *erg.* mechesuárt *gets asphalted* milechesuárt *got asphalted*

b. *impf.* mengesuárt *puts asphalt on* milengesuárt *was putting asphalt on*

c. *pf.* chosuertíi *asphalts it over* chilsuertíi *asphalted it over*

d. *res. state vb.* chelsuárt *asphalted*

e. *antic. state vb.* chesuertáll *is to be asphalted*

(60) PAL n. sébel *shovel* (< Eng. shovel)

a. *erg.* mesébel *gets cleared of debris (by shovelling)* milsébel *got cleared of debris (by shovelling)*

b. *impf.* mesébel[63] *shovels, removes debris from* milsébel *was shovelling/removing debris from*

c. *pf.* sobelíi *shovels it out* silebelíi *shovelled it out*

d. *res. state vb.* selébel *shovelled out, cleared of debris*

e. *antic. state vb.* [none]

(61) PAL n. smengt *cement* (< Eng. cement)

a. *erg.* mesméngt *gets cemented* milsméngt *got cemented*

b. *impf.* mesméngt *puts cement on* milsméngt *was putting cement on*

c. *pf.* simengtíi *cements it over* silemengtíi *cemented it over*

d. *res. state vb.* seleméngt *cemented over*

e. *antic. state vb.* smengtáll *is to be cemented*

(62) PAL n. bomk[64] *pump, small boat engine* (< Eng. pump)

a. *erg.* obómk *gets pumped* ulebómk[65] *got pumped*

b. *impf.* omómk *pumps* ulemómk *was pumping*

c. *pf.* memkíi *pumps it out* milemkíi *pumped it out*

d. *res. state vb.* blomk *pumped out*

e. *antic. state vb.* bemkáll *is to be pumped out*

### 3.2.2.

The great majority of PAL nouns are inflected with *possessor suffixes* to indicate the person and number of a possessor. The categories of possessor expressed in the suffixes correspond exactly to the seven major pronominal distinctions of PAL, as indicated in the list below:

(63)

| Category of Pronominal Possessor | Possessor Suffixes |
|---|---|
| first person singular | -ek, -uk, -ik, -ak |
| second person singular | -em, -um, -im, -am |
| third person singular | -el, -ul, -il, -al |
| first person plural inclusive | -ed, -ud, -id, -ad |
| first person plural exclusive | -am, -(e)mam |
| second person plural | -iu, -(e)miu |
| third person plural (human)[66] | -ir, -(e)rir |

Given the phonetic form of an independently occurring noun stem, it is impossible to predict whether the possessor suffix will take the vowel e, u, i, or a in the three singular forms and in the first person plural inclusive form. This is illustrated by the following examples:

(64)

| Independent Form | | Sample Possessed Form | |
|---|---|---|---|
| reng | *heart, spirit* | rengúk | *my heart, my spirit* |
| ker | *question* | kerím | *your question* |
| oách | *leg* | ochíl | *his/her/its leg, their (non-human) legs* |
| oák | *anchor* | okúd | *our (incl.) anchor* |
| charm | *animal* | chermék | *my animal* |
| chur | *laughter* | cherím | *your laughter* |
| chur | *tongue* | churál | *his/her/its tongue, their (non-human) tongues* |
| chur | *rib of coconut frond* | churúd | *our (incl.) frond* |

As most of the examples (64) also show, PAL noun stems undergo various types of morphophonemic changes when the possessor suffixes are added. These changes affect the vocalic element in the stem and are due mainly to the shift of stress from the noun stem to the possessor suffixes, which are always stressed. Thus, for example, the a of charm *animal* and the u of chur *laughter* are both reduced to a schwa [ə] (orthographic e) in the possessed forms, where they come to appear in unstressed syllables. Similarly, the diphthong represented by oa in oach *leg* and oak *anchor* is reduced to the single vowel o in the corresponding possessed forms.

With regard to the phenomenon of possession, PAL nouns fall into three classes. The examples of (64) are all ***optionally*** possessed nouns because the noun stem can occur as an independent word as well as with possessor suffixes. Bound noun stems that must always have a suffixed pronominal possessor are called ***obligatorily*** possessed nouns — e.g. budek, budem, etc. *my skin, your skin, etc.* or obekuk, obekul, etc. *my older brother, his older brother, etc.* In other words, with obligatorily possessed nouns we do not find independently occurring free forms such as *bud *skin* or *obek *older brother*. Lastly, ***unpossessible*** nouns are those that never take possessor suffixes. A few native PAL nouns and — as we will see in 4.1. below — the majority of loanword nouns belong to this category. In order to show possession with such lexical items, it is necessary to use a prepositional phrase of the form er *of* + ***possessor*** — e.g. delmerab er ngii *his room (lit. room of him)* and sensei er ngak *my teacher (lit. teacher of me)*.

A significant number of loanword nouns have been sufficiently Palauanised so as to participate in the paradigm of noun possession illustrated in (63) above — in other words, they have been borrowed as optionally possessed nouns. Some typical examples are listed below, with only the third person singular form given as a representative example:

(65)

| Loan Source | | PAL Independent Noun | | PAL Possessed Noun | |
|---|---|---|---|---|---|
| Ger. Schraube | *screw* | seráub | | serubél | |
| Jp. taoru | *towel* | táor | | torél | |
| Eng. pipe | | báeb | | bebél | [bɛbɛ́l] |
| box | | baks | | beksél | [bəksɛ́l] |
| (re)bound | | bangd | *rebound, bounce* | bengdél | [bənðɛ́l] |
| dance | | dangs | | dengsél | [ðənsɛ́l] |
| keys | | kíis | | kisél | |
| shovel | | sébel | | sebelél | [səbəlɛ́l] |
| school | | skúul | | skulél | |
| time | | táem | | temél | [tɛmɛ́l] |
| tank | | tangk | | tengkél | [təŋkɛ́l] |

As the examples of (65) clearly show, loanword noun stems undergo the very same phonological processes of vowel and diphthong reduction to which native stems are susceptible.

Many loanword nouns, especially vowel-final stems from Jp., are adopted into the paradigm of noun possession with the addition of a stem-final augment of the form -leng- or -(e)ng-.[67] Note the following examples:

(66)

| | Loan Source | | PAL Independent Noun | PAL Possessed Noun |
|---|---|---|---|---|
| Ger. | Papier | *paper* | babiér | babilngél |
| Jp. | kama | *sickle* | kámang | kamelengél |
| | bando | *belt* | bándo | bandelengél |
| | kata | *shape* | káta | katelengél |
| | skāto | *skirt* | skáto | skatelengél |
| | zubon | *pants* | subóng | subelengél |
| | tama | *ball* | táma | tamelengél |
| Eng. | bucket | | báket | baketengél |
| | bamboo | | bambúu | bambungél |
| | bag | | bek | bekengél |
| | boat | | bos | besengél |
| | book | | buk | bukelengél |
| | cup | | kob | kebengél |
| | matches | | masés | masesengél |
| | soap | | sob | sebengél |

## 3.2.3.

Loanword stems participate rather freely in a large variety of PAL derivational patterns. Perhaps the derivational affix occurring most frequently with foreign noun stems is the prefix ou-, which derives verbs that designate ownership or use of, control over, or participation in the entity designated by the stem itself. First of all, the use of ou- with native stems is illustrated in the examples below:

(67)

| Noun Stem | | Derived Verb in ou- | |
|---|---|---|---|
| blái | *house* | oublái | *own a house* |
| charm | *animal* | ouchárm | *keep a pet* |
| sers | *garden* | ousérs | *keep a garden, do farming* |
| secheléi | *friend* | ousecheléi | *have (particular person) as a friend* |

The great majority of verbs in ou- formed with loanword stems are intransitive verbs derived from nouns. Note the following examples:

(68)

| | Loan Source | | PAL Noun | | PAL Intransitive Verb | |
|---|---|---|---|---|---|---|
| Ger. | Rad | *bicycle* | rrat | | ourrát | *have/own a bicycle* |
| Jp. | bakuchi | *gambling* | bakutsí | *twenty-one (card game)* | oubakutsí | *play twenty-one* |
| | ashi | *foot* | [none][68] | | ouási | *walk/go on foot (rather than ride)* |

| | | | | | |
|---|---|---|---|---|---|
| | denwa | *telephone* | déngua | oudéngua | *have a telephone, make a telephone call* |
| | yakyū | *baseball* | iákiu | ouiákiu | *play baseball* |
| | shibai | *play* | sibái | ousibái | *act in a play* |
| | jidōsha | *auto* | sidósia | ousidósia | *have/own an auto* |
| | zuga | *picture, drawing* | sungá | ousungá | *make a picture/ drawing* |
| Eng. | basket(ball) | | basként | oubasként | *play basketball* |
| | card(s) | | kat | oukát | *play cards* |
| | store | | stoáng | oustoáng | *run/own a store* |
| | tape (recording) | | téib | outéib | *make/have a tape-recording of* |

In a few cases, a borrowed noun prefixed with ou- yields a transitive verb - e.g.

| (69) | Loan Source | | PAL Noun | PAL Transitive Verb | |
|---|---|---|---|---|---|
| Jp. | shirankao | *feigned ignorance* | sirangkáo | ousirangkáo | *pretend that one is innocent* |
| | shiken | *test, exam* | skeng | ouskéng | *examine, give test to* |
| Eng. | dance | | dangs | oudángs | *dance (particular type of dance)* |
| | school | | skúul | ouskúul | *teach* |

Another common PAL prefix is beke-, which is used to form state verbs that designate two divergent and apparently unconnected categories of meaning. First, beke- can be added to noun stems to form state verbs meaning 'prone to..., habitually doing...' Thus, with native stems we have, e.g. beketekoi *talkative* from tekoi *word, speech* and bekesius *swear a lot, talk vividly*, from sius *swearing*. Borrowed noun stems prefixed with beke- and having the connotation of habitual action include PAL mongk *complaint* — bekemongk *always complaining* (cf. Jp. monku *complaint*) and PAL bet *bed* — bekebet *prone to spending a lot of time in bed* (cf. Eng. bed). Second, beke- is also used with noun stems to form state verbs meaning 'having the characteristic smell of...' Native words with this interpretation are, for example, bekecheluch *smell of coconut oil* from cheluch *coconut oil* and bekengikel *smell of fish* from ngikel *fish*. Loanword nouns appearing with the same category of meaning include kosui *perfume* — bekekosui *smell of perfume* (cf. Jp. kōsui *perfume*) and katuu *cat* — bekekatuu *smell of a cat* (cf. Eng. cat).

The process of plural formation in PAL is quite restricted, since the plural prefix re- (r- before a vowel) can only be added to noun stems designating human beings — e.g. chad *person* — rechad *people*, ngalek *child* — rengalek *children*, ekebil *girl* — rekebil *girls*, and so forth. Loanword nouns with human referents can likewise be prefixed with re-, as in sensei *teacher* (< Jp. sensei) — resensei *teachers*, toktang *doctor* (< Eng. doctor) — retoktang *doctors*, etc. In addition, PAL uses a special prefix consisting of the plural morpheme re- followed by another element -chi- to designate nationalities (e.g. Ruk *Truk* — rechiruk *Trukese people*); this prefix, too, can accompany non-indigenous stems, as in Dois *Germany* (< Jp. doisu) — rechidois *Germans* and Merikel *America* (< Eng. American) — rechimerikel *Americans*. Another prefix added to PAL nouns is kl(e)-, which is used to derive abstract nouns meaning something like 'the experience of being...' or 'the act of doing...' Thus, in addition to examples from the native vocabulary such as chad *person, human being*, — klechad *human life, way of life*, we find examples based on loanword stems

such as sensei *teacher* (< Jp. sensei) — klsensei *the experience of being a teacher* and doraib *drive* (< Eng. drive) — kledoraib *the action of driving around in a car*. As one final example of the application of PAL derivational processes to foreign vocabulary, observe how the reciprocal prefix (basic alternant: kau-) can be added to loanword nouns to form related reciprocal verbs — e.g. mondai *problem* (< Jp. mondai) — kaumondai *dispute, argue over* and musing *cooperative enterprise* (< Jp. mujin *mutual financial business*) — kaumusing *participate jointly in a cooperative enterprise*.

## 4. SYNTACTIC PHENOMENA RELATING TO LOANWORDS

By and large, the PAL syntactic system has readily accommodated the various classes of loanwords that have penetrated into the language. Thus, for example, there is no distinction between native and borrowed nouns in terms of their ability to function as sentence subject, sentence object, and so on. In a like fashion, state verbs from non-indigenous sources are virtually indistinguishable from native state verbs with respect to the characteristic syntactic functions of this grammatical class. Nevertheless, there are two areas worthy of mention in which PAL syntax has been especially flexible in dealing with loanwords, one having to do with nouns and the other with verbs.

### 4.1.

As mentioned in 3.2.2. above, PAL nouns fall into three classes with regard to the phenomenon of possession. Thus, in addition to optionally possessed nouns and obligatorily possessed nouns which, respectively, may or must occur with possessor suffixes, PAL has a class of unpossessible nouns that never allow these suffixes. In order to express possession with such nouns, PAL speakers use a periphrastic expression consisting of the preposition er *of* followed by the noun (or noun phrase) indicating the possessor. While there is a small class of native PAL unpossessible nouns consisting primarily of items designating animals, plants, or parts of the natural environment, the great majority of loanword nouns fall into the unpossessible category. As was observed in 3.2.2., a significant, but nevertheless quite small, number of borrowed nouns have been assimilated into PAL to the extent that they participate in the paradigm of noun possession shown in (63) — in other words, they have become optionally possessed nouns. Most borrowed nouns, however, remain unpossessible, and their foreign nature is clearly marked syntactically by the necessity of using a periphrastic expression with them to indicate the possessor. This phenomenon is illustrated below:

(70)

| Loan Source | | | PAL Unpossessible Noun | Sample Use of Unpossessible Noun Followed by Periphrastic Expression of Possession |
|---|---|---|---|---|
| Sp. | rosario | *rosary* | rosário | rosario er a Maria *Mary's rosary* |
| Ger. | Papier | *paper* | babiér | babier er ngii *his paper/letter*[69] |
| | Maschine | *machine, motor* | mesíl | mesil er a dengki *generator* (lit., *machine of electricity*) |

| | | | | |
|---|---|---|---|---|
| Jp. | dōgu | *tool* | dóngu | dongu er a kldaiksang[70] *carpentry tools* |
| | kotai | *answer* | kotái | kotai er a ochur *answer to the math problem* |
| Eng. | clock | | klok *clock, watch* | klok er ngak *my clock/watch* |
| | store | | stoáng | stoa er a Droteo *Droteo's store* |

## 4.2.

As observed at the beginning of 3.1.1., PAL state verbs use the auxiliary mle *was, were* to express the past tense, while action verbs (transitive and intransitive) take the infixed past tense marker -il-. The use of -il- with action verbs of foreign origin is restricted to those few items discussed in 3.2.1. in which a borrowed noun has served as the stem for derived ergative and imperfective forms prefixed with the verb marker me- (or o-) — e.g. imperfective mengiis *open (with a key)* — milengiis *was opening (with a key)* (cf. kiis *key*), etc. Transitive and intransitive action verbs of foreign origin are otherwise impervious to past tense formation with -il-, and instead the native pattern for state verbs is used, since this provides an alternative way (with the auxiliary mle) to mark the past tense overtly. This usage is illustrated in the sentences below: in (71) the borrowed action verb is transitive, while in (72) it is intransitive:

(71) a. v.t. haráu *pay (for)* (< Jp. harau)
A sensei a mle harau a blals.
*teacher* PAST *pay* *fine*
*The teacher paid the fine.*

b. v.t. siraber *examine, investigate* (< Jp. shiraberu)
A bulis a mle siraber er tia el tekoi.
*police* PAST *investigate* *this matter*
*The police investigated this matter.*

c. v.t. skamáer *confront, face, corner* (< Jp. tsukamaeru *catch, seize, arrest*)
Ng techa a mle skamaer er kau?
*it who* PAST *corner* *you*
*Who was it who cornered you?*

(72) a. v.i. bakuháts *explode, blow up* (< Jp. bakuhatsu)
Ng mle bakuhats a tangk.
*it* PAST *explode* *tank*
*The tank exploded.*

b. v.i. séngkio *vote* (< Jp. senkyo [suru])
A betok el chad a mle sengkio.
*many* *person* PAST *vote*
*A lot of people voted.*

c. v.i. saíng *sign name as pledge to abstain from liquor* (< Eng. sign)
A toki a mle saing er a kesus.
PAST *sign* *last* *etc.* *night*
*Toki pledged last night to abstain from liquor.*

State verbs of foreign origin are identical to native state verbs with respect to the use of mle for the past tense. Thus, we have, for example, ksái *foul-smelling* (< Jp. kusai) — mle kusai *was foul-smelling*, kantáng *simple* (< Jp. kantan) — mle kantang *was simple*, and so forth. In addition, PAL state verbs take the auxiliary mo *go, become* (past tense: mlo) to express a change of state — e.g. ungil *good* — mlo ungil *became good, improved*, mekeald *hot* — mlo mekeald *got hot*, etc. The very same pattern is used as well with state verbs of foreign origin, as in sabisí *lonely* (< Jp. sabishii) — mlo sabisi *got lonely* and kitsingái *crazy, insane* (< Jp. kichigai) — mlo kitsingai *went crazy*.

## 5. THE SEMANTICS OF LOANWORDS IN PAL

As we have seen in the examples so far given, the great majority of foreign lexical items have been borrowed into PAL with no discernible semantic modification. There are nevertheless quite a few striking cases in which the original meaning of a loanword has undergone significant distortion. It is difficult to propose a clearcut method of categorising the types of semantic change, but as a rough approximation, the following three-way distinction seems reasonable. First, a process of semantic change called *narrowing* or *specification* involves taking a loanword having a general category (or generic) meaning and using it, in the borrowing language, to designate some specific member, subtype, or subclass of the category in question. Second, a process called *widening* and *expansion* involves just the opposite phenomenon — that is, a term with fairly specialised reference in the contributing language is borrowed with a more general or generic meaning, perhaps being used to refer to the whole class of which the original referent is a part or member. Finally, a third process called *extension* or *shift* does not involve the whole-part or part-whole relationships, respectively, of narrowing or widening, but instead hinges on a part-part relationship. In other words, a given term representing a member of some category is shifted or extended to refer to another comember of the same category, or is simply shifted, somewhat imprecisely, to a referent having similar or related properties.

Semantic narrowing or specification is observed in the loanword items listed below:

(73)

| Loan Source | | | PAL | |
|---|---|---|---|---|
| Ger. | Bild | *picture* | bilt | *holy picture* |
| | Turm | *tower* | turm | *church tower, steeple* |
| Jp. | bangō | *number* | bángngo | *identification number, door number, number in batting order (baseball)* |
| | bokusō | *grass, pasture* | boksó | *elephant grass (used as animal feed)* |
| | ainoko | *half-breed, mulatto* | chainokó | *half-Japanese and half-Palauan* |
| | iroiro | *various, diverse, miscellaneous* | chirochíro | *many-coloured, (children of particular woman) fathered by different men* |
| | ude | *arm* | chúde | *biceps* |
| | eki | *liquid, fluid, solution* | chekí | *battery acid* |

| | | | | |
|---|---|---|---|---|
| | hake | *brush* | háke | *paint brush* |
| | hontō | *main island* | hónto | *Babeldaob* |
| | kankōdan | *tourist group* | kangkódang | *(individual) tourist* |
| | kansoku | *observation, survey* | kánsok | *meteorological survey* |
| | machi | *town, city* | mats | *main town, capital* |
| | minatohan[71] | *harbour quarter* | minatoháng | *area of Koror between Neco Store and T-Dock* |
| | nappa[72] | *greens* | náppa | *cabbage* |
| Eng. | blocks | | blaks | *cement blocks* |
| | ground | | kurángd | *playground* |
| | number | | lámbang | *identification number, telephone number* |
| | sauce | | sos | *soya sauce* |

In some cases, the PAL word not only preserves the original, more general meaning but also develops a more specialised meaning — e.g.

| | | | | |
|---|---|---|---|---|
| (74) Ger. | Gummi | *rubber* | kúmi | *rubber, rubber band* |
| Jp. | bakudan | *bomb* | bakudáng | *bomb, dynamite* |
| | kata | *shape, form* | káta | *shape, form, frame for weaving* |
| Eng. | pipe | | báeb | *pipe, windpipe* |

Semantic widening or expansion appears in loanword items such as the following:

| | | | | |
|---|---|---|---|---|
| (75) Jp. | baikin | *bacillus, bacterium, germ* | baikíng | *disease* |
| | boi | *page, waiter, porter, bellhop* | bói | *servant* |
| | bōkūgō | *air-raid shelter, dugout* | bokungó | *deep hole, pit* |
| | kenpei | *military policeman* | kembéi | *police* |
| | rinbyō | *gonorrhea* | rímbio | *venereal disease* |
| Eng. | rum | | rrom | *liquor* |

In certain cases, the PAL word maintains the original specific meaning and, in addition, develops a more general meaning — e.g.

| | | | | |
|---|---|---|---|---|
| (76) Jp. | otsuri | *change (from purchase)* | otsúri | *change (from purchase), benefit, recompense* |
| | saidā | *cider* | saidáng | *cider, soft-drink* |
| | tanjōbi | *birthday* | tansióbi | *birthday, birthday party* |
| Eng. | buoy | | bói | *buoy, property marker (on land)* |
| | Sunday | | sándei | *Sunday, week* |

The phenomenon of semantic extension or shift is illustrated in the examples below:

| | | | | |
|---|---|---|---|---|
| (77) Ger. | Mark | *mark (monetary unit)* | mak | *fifty cents (old value of one mark)* |
| | schenken | *give, send* | sengk | *gift of money to first-born child by father's family* |
| Jp. | budō | *grape* | búdo | *Panama cherry* |
| | denkibu | *Electric Department* | déngkibu | *power plant* |
| | han | *fief, feudal domain* | hang | *hamlet* |
| | mōchō | *appendix* | mótsio | *appendicitis* |
| | ninjin | *carrot* | ninzín | *type of sweet potato (with orange-coloured inside)* |
| | sarumata | *shorts, trunks* | sarumáta | *panties* |
| | tōgan | *wax gourd, Chinese watermelon* | tongáng | *squash* |
| Eng. | flour | | blauáng | *bread* |
| | ball four | | borhuá | *walk (in baseball)* |

In addition, there are cases in which the PAL word maintains the original meaning and develops one or more coexisting or extended meanings — e.g.

(78)

| | | | | |
|---|---|---|---|---|
| Ger. | Papier | *paper, document* | babiér | *paper, document, letter, book* |
| Jp. | tama | *ball, marble* | táma | *ball, marble, fried flour ball* |
| Eng. | cat | | katúu | *cat, sweetheart* |
| | clock | | klok | *clock, watch* |
| | soursop | | sáusab | *soursop, spade or heart in cards*[73] |

## ACKNOWLEDGEMENTS

I would like to thank Romana Anastacio for her invaluable assistance in providing me with information on various aspects of Palauan loanword vocabulary. Appreciation is also due to Wako Hinds and to Charles DeWolf, who provided me with helpful insights on Japanese and Palauan, respectively. Theo Faustino kindly suggested many revisions to the original draft.

## NOTES

1. The Palau Islands were discovered in 1543 by the Spanish explorer Ruy López de Villalobos, who named them the Arrecifos. It is said that Sir Francis Drake visited the islands in 1579.

2. For a more detailed analysis of the PAL phonological system, see Josephs 1975: Chapter 1.

3. For Sp., Ger., and Eng., the loan source is given in the modern standard orthography. For Jp., the Romanisation system used is that found in Kenkyûsha's *New Japanese-English dictionary,* ed. Koh Masuda (Fourth edition, Tokyo, 1974). This system presents a phonetic (rather than phonemic) spelling of Jp. syllables using Eng. orthographic equivalents for the sounds in question. Thus, the syllables of Jp. are rendered as follows:

| a | i | u | e | o |
|---|---|---|---|---|
| ka | ki | ku | ke | ko |
| sa | shi [ši] | su | se | so |
| ta | chi [či] | tsu [cu] | te | to |
| na | ni | nu | ne | no |
| ha | hi | hu [Φu] | he | ho |
| ma | mi | mu | me | mo |
| ya | | yu | | yo |
| ra | ri | ru | re | ro |
| wa | | | | |
| ga | gi | gu | ge | go |
| za | ji [ǰi] | zu [ʒu] | ze | zo |
| da | | | de | do |
| ba | bi | bu | be | bo |
| pa | pi | pu | pe | po |

In addition to the above, a *palatal glide* can occur between a consonant and a following a, u, or o, as indicated in the following syllables:

| | | |
|---|---|---|
| kya | kyu | kyo |
| sha [ša] | shu [šu] | sho [šo] |
| cha [ča] | chu [ču] | cho [čo] |
| nya | nyu | nyo |
| hya | hyu | hyo |
| mya | myu | myo |
| rya | ryu | ryo |
| gya | gyu | gyo |
| ja [ǰa] | ju [ǰu] | jo [ǰo] |
| bya | byu | byo |
| pya | pyu | pyo |

Long vowels within a single syllable are indicated by a long mark (e.g. kū, gō, etc.), and geminate consonants are spelled double (e.g. tt, pp, etc.). The symbol n stands for a syllable-final nasal phoneme that assimilates to the initial consonant of the following syllable within the word (e.g. shinpai [šimpay] *worry*, kankei [kaŋkey] *relationship*) and is realised in word-final position as a nasalised version of the preceding vowel (e.g. hon [hoõ] *book*, shinbun [šimbuũ] *newspaper*). In Tokyo dialect, the g in syllable-initial position is usually pronounced [ŋ] word-internally (e.g. negi [nɛŋi] *green onion*). The location of the pitch accent, although marked in Kenkyûsha's Romanisation, is omitted here.

4. The orthographic conventions of PAL should become clear from the ensuing discussions, but note the following equivalents:

| PAL orthography | Phonetic equivalent |
|---|---|
| ch | [ʔ] |
| ng | [n, ŋ] |
| d | [ð, θ] |
| ll | [l:] |
| rr | [r̃] |

In this paper, the primarily stressed syllable of a polysyllabic word is marked with an accent over the vowel or the peak of a diphthong, although no such marking is made in standard PAL orthography. Note, further, that if no PAL gloss is given, the word in question has the same meaning as in the contributing language.

5. Some minor details of pronunciation in the contributing languages are disregarded here. For example, whereas orthographic b and v of Sp. are both pronounced as bilabial stops in word-initial position, intervocalically they are normally articulated as lenis bilabial fricatives. For further details on the Sp. situation, see Stockwell and Bowen 1965:47-48.

6. In the PAL form, the final l of Sp. barril has been idiosyncratically lengthened.

7. Note that the Sp. plural form cebollas was reinterpreted in PAL as a singular.

8. In Ger. orthography, w stands for [v].

9. In this word, the final nasal is pronounced [n] rather than [ŋ] — hence, the unusual spelling with n rather than ng.

10. Indeed, the irregular spelling of PAL Papa with p's is intended to highlight this exception.

11. For a detailed discussion of PAL a, see Josephs 1975: Section 2.6. When adopting certain PAL nouns and place names, Eng. speakers also had trouble with PAL a, interpreting it as an integral part of the stems in question. Thus, in Eng. texts, we find the PAL community house bai being referred to as an abai and the villages of Irrai and Imeliik spelled as Airai and Aimeliik.

12. In Jp. itself, /p/ is a rather restricted phoneme, occurring only geminate in words of native origin. Otherwise, the occurrence of /p/ is limited to borrowings from Chinese and Eng. and to a special class of 'onomatopoeic' vocabulary. See McCawley 1968:77ff. for further details.

13. Jp. kyabu (longer alternant: kyaburetā) is itself a borrowing from Eng. A significant number of PAL loanword items can be traced through Jp. (from which they were directly borrowed) to their ultimate Eng. source. Thus, in certain cases, characteristic distortions of the original Eng. pronunciation clearly suggest direct borrowing from Jp. — e.g. Jp. and PAL bando *belt* (< Eng. band); Jp. puragu *plug* — PAL berangu *spark plug* (< Eng. plug); Jp. and PAL kanaria *canary* (< Eng. canary); Jp. and PAL masku *mask* (< Eng. mask); Jp. merikenko *flour* (< Eng. American + Jp. ko *powder*) — PAL merikengko *flour;* Jp. ranningu *athletic shirt* — PAL ranningngu *athletic shirt* (< Eng. running [shirt]); Jp. rajiētā *radiator* — PAL razieta *radiator* (< Eng. radiator). In other cases, however, it is very difficult to tell whether a given loanword was taken directly from Jp. or from its original Eng. source — e.g. PAL batteri (cf. Eng. battery, Jp. batterī), PAL chea (cf. Eng. air, Jp. ea), PAL dainamo (cf. Eng. dynamo, Jp. dainamo), PAL saireng (cf. Eng. siren, Jp. sairen).

14. In some dialects of Jp. (e.g. downtown Tokyo), /hi/ is pronounced [ši] rather than [çi].

15. Many speakers of the older generation, who still have knowledge of Jp. and spoke the language fluently during the period of the mandate, tend to maintain the original Jp. pronunciation for the allophones of /h/. Thus, for example, the initial fricative of PAL hútsu *common, usual* (cf. Jp. futsū) will often be heard as [Φ] or [f] rather than [h].

16. In present-day PAL, [x] appears in the words of (9) and sporadically among younger speakers as a replacement for the [$k^h$] allophone of /k/ when the sound in question occurs word-finally after a stressed vowel Thus, for such speakers, chermék *my animal* is pronounced as [ʔərméx] rather than as [ʔərmék$^h$]. Interestingly enough, it is reported that as late as the Ger. times, many PAL speakers (at least the older generation) pronounced the present-day phoneme /ʔ/ as [x]. Following Ger. orthographic conventions, the letters ch were used to spell this sound (cf. Ger. doch [dɔx] *therefore*), and this spelling has remained standard even after the phonetic change from [x] to modern-day [ʔ]. Only one loanword with [h] appears to have been taken from Ger. — namely, the PAL interjection hall *stop!* (cf. Ger. Halt! *stop!*). The origin of the unusual h-initial pair hngong (exclamation to draw attention to a pleasant smell) and hngob (exclamation to draw attention to an unpleasant smell) is obscure.

17. Intervocalically, Sp. /d/ is actually pronounced as a lenis voiced fricative similar to [ð]. Cf. note 5 and see Stockwell and Bowen 1965:44-47.

18. The loss in PAL of the second d of soldado is probably due to the fact that the Sp. sound is so lenis (cf. note 17) as to appear nearly inaudible to non-native speakers. This phenomenon is mentioned by Stockwell and Bowen 1965.

19. In Ger. (morphophonemically basic) voiced stops are automatically devoiced in word-final position. Thus, the /d/ at the end of Rad was heard as [t] and borrowed into PAL as /t/ (realised word-finally as [$t^h$]). The same phenomenon is observed for Ger. Bild [bilt] *picture* — PAL bilt [$bilt^h$] *holy picture*.

20. The sound represented by word-final orthographic z in Sp. is either [s] or [θ].

21. Sp. orthographic 'soft' c (i.e. c before i or e) is pronounced either as [s] or [θ]. Cf. note 20.

22. In all of these examples, Eng. /z/ [z] is represented by orthographic s. PAL kíis is used as a singular although borrowed from the Eng. plural keys (cf. note 7).

23. In this and the following example, a schwa [ə] (orthographic e) has been inserted to avoid the prohibited consonant cluster sr.

24. In Jp. shiken, the high front vowel i is voiceless between two voiceless consonants — i.e. we have [šḭkɛẽ]. In PAL, the voiceless vowel [ḭ] is lost completely — hence, skeng [skɛŋ]. For more discussion of this phenomenon see 2.2.3. below.

25. Note the unexpected optional i in the PAL form, indicating that the Jp. word might have been misinterpreted as *uchusu by some speakers.

26. A good number of younger speakers with a knowledge of Eng. pronounce these words with initial [t͡š] rather than [t͡s]

27. Because both the [t͡š] of inch [Int͡š] and the [d͡ž] of hinge [hInd͡ž] become PAL s and the h of Eng. hinge is lost, the PAL form íings is homophonous.

28. Note the similarity to the phoneme /t/, which also has an aspirated allophone [$t^h$] word-finally (cf. 2.1.2.).

29. Sp. orthographic ñ represents a palatised nasal — hence, the appearance of i in the PAL counterpart.

30. Word-final PAL ng will be discussed presently.

31. Even a few Eng. borrowings show the change of n to l — e.g. PAL lámbang (< Eng. number), kombalíi (< Eng. company), and Lukiléi (< Eng. New Guinea).

32. In this interesting example, PAL has [ŋ] before [k] even though the Eng. source does not have [ŋ] (due, undoubtedly, to the morpheme boundary intervening between the final nasal of sun and the initial velar of glasses). Note, in addition, that in the PAL form, the Eng. plural suffix has been omitted.

33. Interestingly enough, no examples were observed of Eng. loanwords with word-final [ŋ]. One or two exceptions exist to the pattern observed here — e.g. PAL taifún (< Eng. typhoon), where word-final Eng. [n] = PAL [n].

34. One notable exception to this statement is PAL kámang *(arm) twisted/ crippled* (< Jp. kama *sickle, hook*).

35. Note, in addition, the existence of a rule in the native phonology that requires the assimilation of /l/ in certain morphemes to an /r/ elsewhere in the word. Thus, while the basic form of the past tense infix is /-il-/, as in milleseb *burned* (cf. meleseb *burn*) and chilarm *suffered* (cf. chuarm *suffer*), the l changes to r in such forms as mirrael *travelled* (cf. merael *travel*) and riros *drowned* (cf. remos *drown*).

36. In this example, the a of azucár has evidently been analysed as the PAL prenominal particle a, thus leaving -zucár (= PAL sukál) as the independent noun (cf. note 11).

37. Babilngél *his paper/letter/book* is the third person singular possessed form of babiér. Note the change of r to l in the possessed form and see 3.2.2. below for further discussion.

38. The form kil is also used by some speakers.

39. Although the occurrence of stress is probably predictable within the native phonology (see Wilson 1972:39-42), there do not appear to be any consistent patterns governing stress in loanword phonology. Whether the contributing language has a stress accent system (as in Sp., Ger., and Eng.) or a pitch accent system (as in Jp.) does not seem to matter: it is virtually impossible to predict which syllable of the PAL loanword will receive primary stress. Clearly, much more research is needed in this challenging area.

40. The tense vowels in these pairs are sometimes analysed phonetically as diphthongs — i.e. [iy], [uw], and [ey].

41. There are some further conditioning factors that determine whether or not a given high vowel will devoice (e.g. the location of the pitch accent within the word and the number of devoiceable vowels potentially occurring in adjacent syllables), but these are not relevant to the present discussion. For a complete analysis of Jp. vowel devoicing, see Han 1962:17-45.

42. In this word, Jp. h before i [ç i] becomes s in PAL with concomitant loss of the vowel. Similar examples are Pal. skózio (< Jp. hikōjō *airport*) and PAL skidás (< Jp. hikidashi *drawer*).

43. With the loss of devoiced u, the PAL form gets an initial cluster tsk-, which is optionally simplified to sk-. Another similar example is PAL (t)skáreter (< Jp. tsukarete iru *tired*).

44. In this word, Jp. ch [t͡ɕ] gives PAL ts; and after the loss of the word-final devoiced vowel, the consonant cluster nts simplifies to ns.

45. Alternate pronunciations, including the glide articulation, are [kmɛ:ʏð] [ði:ʏl] [ðəgó:ʍl], and [bu:ʍʔ]

46. For a more detailed discussion of the syntactic and semantic properties of PAL state verbs, see Josephs 1975: Chapter 7.

47. This is a slang form for standard asemo.

48. This PAL state verb illustrates a humorous semantic shift from the original meaning of the Eng. noun.

49. In comparing Eng. **American** with PAL Merikel, we again observe the phenomenon in which an initial a- in the source word has been interpreted as the prenominal particle a, thus leaving -merican as the basis for the independent PAL noun. Cf. notes 11 and 36.

50. This word is also used in PAL as a noun meaning *joke*.

51. Note the unpredictable meaning distortion here, possibly influenced by contamination from daiziob, which has the same meaning and is partially similar phonetically.

52. The Jp. negative equivalent of atte iru — namely, awanai *not suitable/appropriate (for each other)* — is also borrowed into PAL as chauanái.

53. This word is also used as an intransitive verb in PAL — i.e. it can also have the meaning *go flat, get punctured, blow out.*

54. The PAL forms can also be used as nouns with the meanings, respectively, *(a) promise, (b) inspection, examination.*

55. PAL saíng (note the unusual accentuation) is also used as a noun meaning *sign* and as an intransitive verb meaning *sign name (especially as pledge to abstain from liquor).*

56. PAL kats is also used as a noun meaning *winner.*

57. PAL máke is also used as a noun meaning *loser.*

58. Jp. mawaru is the intransitive partner of mawasu (cf. (55) above).

59. The PAL form changár is also used as an intransitive verb meaning *(salary, etc.) increase, (person) get excited/nervous.*

60. PAL has also borrowed the negative equivalent kingatskanái *not be aware of, not notice* (< Jp. ki ga tsukanai).

61. The many details omitted here can be found in Josephs 1975: Chapters 5-7, 9-12.

62. In other words, the basic form of imperfective mengélebed is (verb marker) me- + (imperfective marker) -ng- + (stem) chelébed, and by a special morphological rule, the initial ch of the stem is deleted directly following the imperfective marker.

63. In the imperfective forms of some verbs with s-initial stems, the imperfective marker is idiosyncratically absent, thus making the imperfective forms homophonous with the ergative forms. This phenomenon is also observed in example (61).

64. Note the unusual dissimilation of word-final b to k in this form. An alternant without dissimilation — i.e. bomb — is also in use.

65. In verbs with o- as the verb marker, the past tense is formed by replacing o- with ule-. The sequence ule- may be a phonologically changed form of (verb marker) o- + (past tense marker) -il- in which the contiguous vowels mid back o and high front i have blended to yield high back u. See Josephs 1975:6.3.2. for more discussion of the phenomenon of vowel blending in PAL.

66. The suffixes -ir and -(e)rir designate *human* plural possessors only. In order to express a non-human plural possessor, the third person singular suffixes -el/-ul/-il/-al are required, thus resulting in some potential ambiguity.

67. The -ng- augment is also found in the possessed forms of certain native stems — e.g. bas *charcoal* — besengél, bilás *boat* — bilsengél, úum *kitchen* — umengél, and téu *width* — tengél.

68. Interestingly enough, there is no independent PAL noun asi *foot* but only the derivative ouasi.

69. The borrowed noun babiér *paper, letter* is 'semi-possessible' because some speakers use possessor suffixes (e.g. babilngél *his paper/letter* of (66)) while others indicate possession with a periphrastic expression, as given here.

70. Kldaiksang *being a carpenter* is an abstract noun formed by adding the prefix kl(e)- (cf. the last paragraph of 3.2.3. above) to the borrowed noun daiksang *carpenter* (< Jp. daiku *carpenter* + san [title of address]).

71. This term is archaic in Jp.

72. This is a Jp. dialect word.

73. The meaning shift here is based on an analogy between the shape of the soursop fruit and that of the heart and spade symbols in cards. To distinguish hearts from spades the following terms are used: bekerekard el sausab *heart (lit. red soursop)* and chedelekelek el sausab *spade (lit. black soursop)*.

## BIBLIOGRAPHY

HAN, Mieko Shimizu

1962 *Japanese phonology: an analysis based upon sound spectrograms.* Tokyo: Kenkyusha.

JOSEPHS, Lewis S.

1975 *Palauan reference grammar.* Honolulu: University Press of Hawaii.

McCAWLEY, James D.

1968 *The phonological component of a grammar of Japanese.* The Hague: Mouton.

McMANUS, Edwin G., S.J.

1977 *Palauan-English dictionary,* ed. and expanded by Lewis S. Josephs, with the assistance of Masa-aki Emesiochel. Honolulu: University Press of Hawaii.

STOCKWELL, Robert P. and J. Donald BOWEN

1965 *The sounds of English and Spanish.* Chicago: University of Chicago Press.

WILSON, Helen Irene

1972 *The phonology and syntax of Palauan verb affixes.* (Doctoral dissertation, Department of Linguistics, University of Hawaii, Honolulu.) *Working Papers in Linguistics, University of Hawaii* 4/5.

# COMPLEMENTATION IN PALAUAN

Lewis S. Josephs

## 1. INTRODUCTION

A survey of extant works on Palauan (PAL) reveals a steadily expanding interest in elucidating the syntactic properties of the language, and in the post-war period especially, much emphasis has been placed on the investigation of complex sentences, in particular those involving various types of subordination. In this essay, a specific type of subordination — namely, complementation — will be examined in detail, and contrasting complement structures will be identified in terms of salient syntactic and semantic characteristics.

## 2. CHARACTERISTICS OF OBJECT COMPLEMENTS

In PAL, as in all other languages, whole sentences (or propositions) can function as the subjects or objects of verbs. Thus, embedded sentences dominated by the subject NP (noun phrase) node are called ***subject complements***, while those dominated by the object NP node are termed ***object complements***. We shall focus first on the latter type.

There is an important class of PAL transitive verbs that can take either concrete objects or objects that describe actions or activities. These two contrasting object types are illustrated in the examples below:

(1) a. Ak mla mo merek er a subelęk.
*I have finish my homework*
*I have finished my homework.*

b. Ak mla mo merek el meruul a kall.
*I have finish prepare food*
*I have finished preparing the food.*

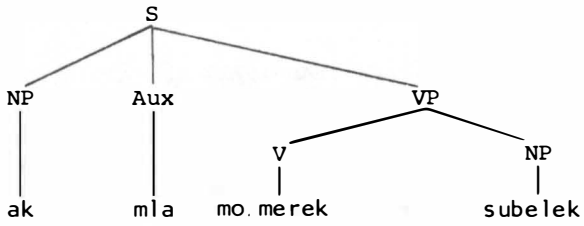
(2) a. A Droteo a milsuub er a reksi.
*learned history*
*Droteo learned history.*

b. A Droteo a milsuub el meruul a mlai.
*learned make canoe*
*Droteo learned how to make canoes.*

Byron W. Bender, ed. ***Studies in Micronesian Linguistics***, 125-148. ***Pacific Linguistics***, C-80, 1984.

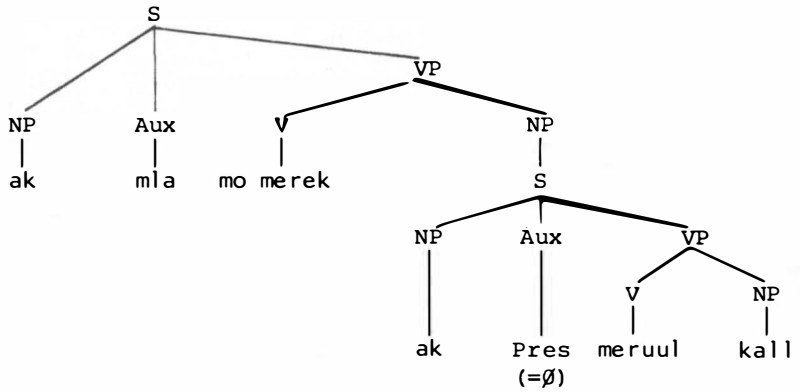
In the (a)-sentences above, the objects of mo merek *finish* and mesuub *learn, study* are the single concrete nouns subelek *my homework* and reksi *history*,[1] while in the (b)-sentences the objects of these same verbs designate certain actions or activities (preparing food, making canoes). Since the object in the (b)-sentences contains a finite verb but no overtly-expressed subject, we can conclude that it is ultimately derived from a fully specified embedded sentence whose subject has been deleted under identity with that of the 'higher', main clause. Thus, while 1a, for example, has the following deep structure (simplified with respect to irrelevant details),

(3)




sentence 1b has the more complex structure below:

(4)




In the deep structure configuration 4, the expansion of the object NP into a fully specified sentence (the object complement) reflects the semantic fact that in 1b, the object of mo merek *finish* is an entire activity (preparing food) performed by a particular agent (ak *I*). The semantic interpretation of 1b requires a deep structure in which the agent (= subject) of the embedded proposition is identical to that of the higher, main clause, since 1b can only be understood to mean that the person preparing the food and the person completing that preparation are one and the same individual. For this reason, the verb mo merek is said to require the *Like-Subject Constraint*: in deep structures like 4, the subject of the object complement must always have the same referent as that of the verb mo merek itself. In deriving the actual surface structure 1b, the complement subject ak is deleted by a process called *Equi-NP Deletion*, and an insertion rule introduces el at the beginning of the complement. As we will see, el marks many varieties of complements in PAL and is thus justifiably labelled a *complementiser*.[2]

In addition to having no overtly expressed subjects, the object complements of 1b and 2b require verb forms in the present tense, even when the verb of the main clause designates some past tense — e.g. mla mo merek *has finished* (recent

past), milsuub *learned* (non-recent past) — and the whole sentence therefore describes a past action or event. Thus, sentences like the following are ungrammatical because the (italicised) verb of the object complement is in some past tense:

(1b') *Ak mla mo merek el {*mirruul* / *mla meruul*} a kall.

*I have finish {prepared / have prepared} food.*

(2b') *A droteo a milsuub el {*mirruul* / *mla meruul*} a mlai.

*learned {made / have made} canoe*

The apparent 'neutralisation' of tense marking in the embedded sentences of 1b and 2b is a characteristic property of certain types of complement structures[3] as opposed to, say, relative clauses. Thus, as the following example shows, there are no restrictions whatsoever on the tense markers which can occur in PAL relative clauses:

(5)

A chad el { a. mo meruul / b. meruul / c. mla meruul / d. mirruul } a mlai a Sebastian.

*person* { a. *will make* / b. *makes* / c. *has made* / d. *made* } *canoe*

*The person who* { a. *will make* / b. *makes* / c. *has made* / d. *made* } *the canoes is Sebastian.*

The differences in freedom of occurrence of tense markers in complement structures as opposed to other structures (e.g. relative clauses) may well represent a universal phenomenon. Thus, Kajita 1967:47-58 notes that in Eng., the object complements accompanying verbs like *hesitate, fail, volunteer,* etc. are restricted in their internal structure such that only the present tense form of the infinitive can occur:

(6) a. John hesitated to go/*have gone.
b. Bill failed to inform/*have informed the teacher.
c. Mary volunteered to drive/*have driven the limosine.

The starred verb phrases in 6 show that the object complement cannot contain past (or perfective) markers. By contrast, Eng. relative clauses have no tense restrictions, as the translation in 5 clearly illustrates.

Returning to the deep structure configuration of 4, we can now see that the embedded sentence which represents the object complement cannot be expanded with total freedom as can the highest (matrix) sentence. First of all, the NP subject of this complement cannot be *any* noun or pronoun, but must correspond to the subject of the matrix clause; and, second, the Aux node of the complement cannot be *any* tense marker at all, but must be specifically expanded as Pres (present tense).

## 2.1. Further examples of object complements

In the sentences below we observe further instances of object complements following the transitive verb mo merek *finish*:

(7) a. Ak mlo merek el remurt er a eai el klok.
*I finished rwn at eight clock*
*I finished running at eight o'clock.*

b. A Droteo a mla mo merek el bechiil.[4]
*has finished married*
*Droteo is no longer married.*

c. Ke mla mo merek el omengur?
*you have finished have dinner*
*Have you finished eating?*

d. Ke mla mo merek el melamech a dekool?[5]
*you have finished smoke cigarette*
*Have you finished smoking your cigarette?*

Other transitive action verbs such as omuchel *begin* and melasem *try (out)* resemble mo merek in allowing both concrete objects (as in the (a)-sentences) and object complements (as in the (b)-sentences):

(8) a. Kede mo omuchel er a blai er a klukuk.
*we will begin house tomorrow*
*We will begin (to build) the house tomorrow.*

b. Te ulemuchel el meruul er a blai er a elii.
*they began make, build house yesterday*
*They began to build the house yesterday.*

(9) a. Ke mla melasem er a sasimi?[6]
*you have try out sashimi*
*Have you tried out/tasted the sashimi?*

b. Ak millasem el menga er a ngikel.
*I tried eat fish*
*I tried to eat the fish.*

PAL has a small class of transitive state verbs[7] that can be identified by the following two features: first, like all other transitive verbs, they can take objects; and, second, like all other state verbs, they have past tense forms derived with the auxiliary mle *was, were*. Two transitive state verbs — meduch *know how (to), be skilled at* and metitur *be ignorant of, not know how (to), not be capable of* — can be followed by concrete objects as well as object complements. Thus, the examples below parallel those of 8 and 9 above:

(10) a. Ak mle meduch er a ochur.
*I was skilled at math*
*I used to be good at math.*

b. A Droteo a mle meduch el melekoi a tekoi er a Siabal.
*was skilled at speak Japanese*
*Droteo used to be good at speaking Japanese.*

(11) a. Ak metitur er a mesil er a sidosia.
*I am ignorant of motor of car*
*I don't know anything about (fixing) car motors.*

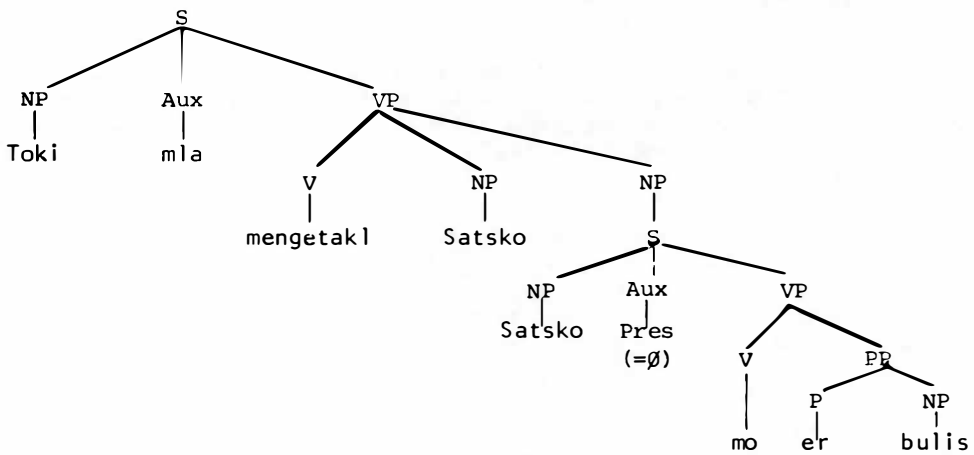
b. A John a metitur el mengikai.
*is ignorant of swim*
*John doesn't know how to swim.*

In all of the examples given so far, the unexpressed subject of the object complement is understood to be identical to that of the preceding higher verb. There are some cases, however, in which the unexpressed subject of the object complement is understood differently, as in the sentence below:

(12) A Toki a mla mengetakl er a Satsko el mo er a bulis.
*has persuaded go to police*
*Toki has persuaded Satsko to go to the police.*

Here, the higher subject (Toki) influenced some other person (Satsko) to do a particular action (i.e. go to the police). In other words, the action of going to the police was carried out by the person persuaded (Satsko) rather than by the person doing the persuading (Toki). Therefore, the unexpressed subject of the object complement el mo er a bulis *(to) go to the police* cannot be identical to Toki, the subject of mla mengetakl *has persuaded*, but rather must be identical to Satsko, the noun immediately following mla mengetakl. In 12, then, the verbal expression mla mengetakl appears to be followed by a succession of two objects, the first a noun identifying the person influenced or affected by the persuasion and the second an object complement describing the action pursued as a result of the persuasion. This two-object structure, in which the first object may be termed 'indirect' and the second 'direct', is illustrated in 13, the deep structure of 12:

(13)




In the derivation of 12 from 13, the identical subject Satsko of the object complement is deleted under identity with the indirect object Satsko of the main clause, and the complementiser el is inserted in complement-initial position.

The two-object structure described above is typical of other PAL verbs such as olengeseu *help*, olisechakl *teach*, and oldurech *tell, ask*. Observe the examples below:

(14) a. Ak ullengeseu er a Toki el meruul a subelel.
*I helped do her homework*
*I helped Toki do her homework.*

b. Ak ngilsuterir[8] a resechelik el mengetmokl er a blai.
*I helped them my friends clean house*
*I helped my friends clean the house.*

c. A rubak a ullisechakl er a Droteo el melasech a mlai.
*old man taught carve canoe*
*The old man taught Droteo how to carve canoes.*

d. A sensei a ulderchak[9] el mo er a Guam.
*teacher told me go to*
*The teacher told me to go to Guam.*

## 3. CHARACTERISTICS OF SUBJECT COMPLEMENTS

Any adequate description of PAL syntax must recognise a process of ***subject extraposing,*** as illustrated in the pairs of sentences below:

(15) a. A sechelik a mla mei.
*my friend has come*
*My friend has come.*

b. Ng mla me a sechelik.
*he has come my friend*

(16) a. A biang a mekelekolt.
*beer cold*
*The beer is cold.*

b. Ng mekelekolt a biang.
*it cold beer*

To PAL speakers, the (a) and (b)-sentences in 15-16 are equally natural and acceptable. While the (a)-sentences show the normal word order ***subject — predicate,*** in the (b)-sentences the subjects (sechelik *my friend* and biang *beer*) have been shifted, or ***extraposed,*** to the right of the predicates (mla mei *has come* and mekelekolt *is cold*). Thus, the (b)-sentences are derived from the (a)-sentences by a process of subject extraposing, which leaves in subject position a ***pronominal trace*** of the extraposed subject. Since the extraposed subjects of 15-16 are third person singular nouns, the third person singular non-emphatic pronoun ng *he, she, it* remains as a trace. If, however, the extraposed subject is third person (human) plural, the pronominal trace must be te *they*, the corresponding third person (human) plural non-emphatic pronoun. With 15 above, contrast the following pair:

(17) a. A resechelik a mla mei.
*my friends have come*
*My friends have come.*

b. Te mla me a resechelik.
*they have come my friends*

The only difference between 15 and 17 is that the subjects are singular (sechelik *my friend*) and plural (resechelik[10] *my friends*), respectively; in the former case the pronominal trace is ng, while in the latter it is te.

PAL has a special group of obligatorily possessed nouns[11] which, among other things, refer to such concepts as liking or disliking, as illustrated in the examples below:

(18) a. Ng soak a biang *I like beer./I'd like some beer.*
*my liking, beer*
*what I like*

b. Ng sorir a buuch. *They like betel nut./They'd like some betel nut.*
*their liking, betel*
*what they like nut*

(19) a. Ng chetik a milk. *I dislike milk./I don't want any milk.*
*my disliking,*
*what I dislike*

b. Ng chetil a dekool. *He dislikes cigarettes./He doesn't want any cigarettes.*
*his disliking, ciga-*
*what he dislikes rette*

Although the Eng. equivalents for the above examples include verbs (*like, dislike*), the PAL sentences do not contain any verbs at all but instead have obligatorily possessed nouns such as soak *my liking, what I like* and chetil *his disliking, what he dislikes* followed by concrete nouns such as biang *beer* and dekool *cigarette*. Therefore, these sentences appear to mean, literally, *My liking is beer, His disliking is cigarettes,* and so on.

What is striking about all of the sentences in 18-19 is the appearance of ng in initial (subject) position. The only plausible explanation for this ng is that it has arisen as a pronominal trace of an extraposed subject. Thus, we propose that a sentence like 18a is derived from the following source by a process of extraposition:

(20) A biang a soak. *Beer is what I like.*

Example 20 is a copular sentence in which two noun phrases (biang *beer* and soak *my liking, what I like*) are being equated. The structure of 20 is the same as that of copular sentences such as the following:[12]

(21) A Droteo a sensei er a ochur.
*teacher of math*
*Droteo is a math teacher.*

Interestingly enough, copular sentences like 20 are not used very frequently by speakers of PAL. Such a sentence would be uttered only for the purpose of imparting a special emphatic connotation — i.e. 20 implies that it is beer and only beer (out of a choice of various possible beverages) that the speaker has come to like. Under normal circumstances, a copular sentence like 20 — namely, one containing as its second element a form of the obligatorily possessed nouns soal or chetil — must undergo the process of subject extraposing to yield sentences like 18-19. In other words, while subject extraposing is optional with sentence types such as those of 15-16 (in which both the (a) and (b)-alternatives are equally acceptable), this process is nearly *obligatory* with copular sentences containing soal and chetil.[13]

Now, with 18a and 19a compare the following sentences:

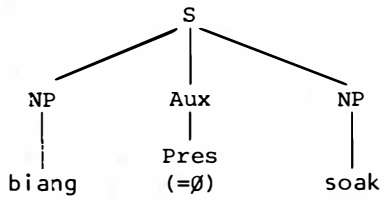
(22) a. Ng soak el melim a biang
*my liking, drink beer*
*what I want*
*I want to drink some beer.*

b. Ng chetik el melim a milk.
*my disliking, drink*
*what I don't want*
*I don't want to drink any milk.*

Whereas the sentence-final elements in 18a and 19a are extraposed subject noun phrases (biang and milk), in 22a-b the sentence-final elements el melim a biang and el melim a milk are reduced clauses containing a finite verb (melim *drink*) but no overtly expressed subject. Although the subject of these clauses is not expressed, it is nevertheless understood to be identical to the pronominal possessor 'my' designated by the suffix (-ak or -ik) on the preceding obligatorily possessed noun (soak or chetik).

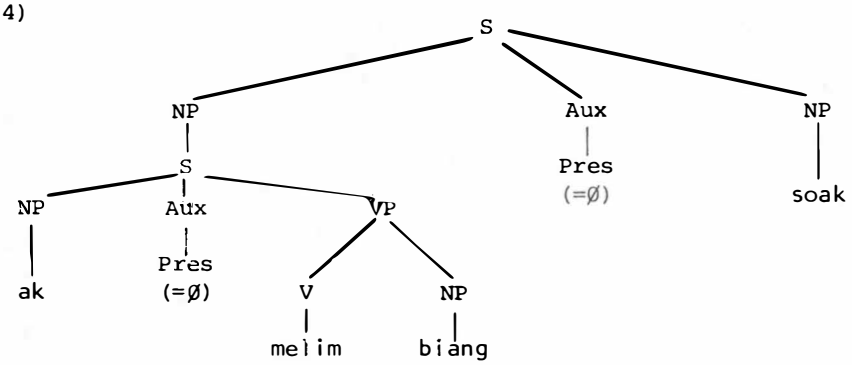
The abovementioned phenomena can be accounted for if we assume that sentences like 22a-b have as their subjects not simple noun phrases, but entire sentences representing activities performed by a specific agent. Thus, while 18a is derived by subject extraposing from a relatively simple copular structure of the form

(23)




a sentence like 22a has its source in the following more complex structure:

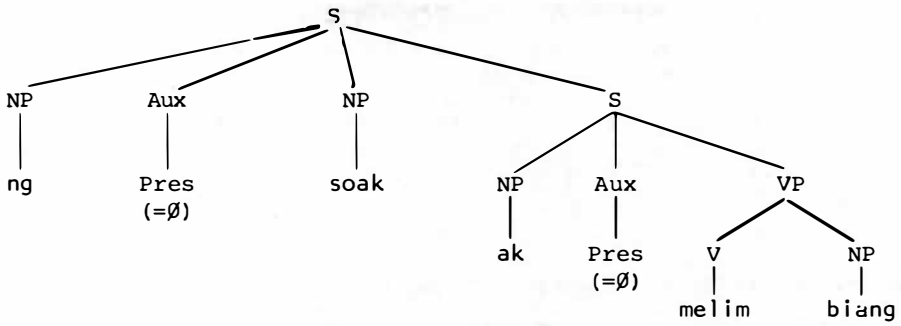
(24)




Although 24 is a copular sentence like 23, its subject has the structure of an entire sentence — i.e. the subject NP of 24 dominates a ***subject complement***. Just as in the case of object complements, the structure of the subject complement in 24 is restricted in two significant ways: first, the embedded subject (ak *I*) cannot be freely chosen but must correspond to the pronominal possessor of the obligatorily possessed noun soak in the matrix sentence;[14] and, second, the verb of the subject complement must be in the present tense.

In deriving 22a from 24, the subject complement is obligatorily extraposed, leaving the pronominal trace ng.[15] Thus, the following intermediate structure results:

(25)




Then, the subject ak *I* of the extraposed subject complement is deleted under identity with the possessor suffix -ak of soak and the complementiser el is inserted at the beginning of the complement.

From the above discussion we can see that PAL subject complements and object complements are restricted in identical ways and must undergo identical transformations such as Equi-NP deletion and complementiser insertion. Furthermore, it is clear that subject extraposing is obligatory in copular sentences with soak and chetik, etc. regardless of whether the deep structure subject is a simple noun phrase (as in 18-19) or a fully specified embedded sentence — i.e. a subject complement (as in 22). It should be noted in passing that deep structures like 23 semantically represent either *general statements* or statements about someone's desire (or lack of desire) on a *specific occasion*, while deep structures like 24 can only represent specific occasions. Because 19a, repeated here for convenience,

(19a) Ng chetik a milk. *I dislike milk./I don't want any milk.*
*my disliking, what I dislike*

is derived from a structure like 23, chetik refers either to a general, habitual distaste (chetik = *I dislike* [*on all occasions*]) or to a lack of desire on a specific occasion (chetik = *I don't want* [*on this particular occasion*]). By contrast, 22b

(22b) Ng chetik el melim a milk. *I don't want to drink any milk.*

can only constitute a refusal on a specific occasion.[16]

As we have seen above, the obligatorily possessed nouns soal *his/her liking* and chetil *his/her disliking* can have either simple noun phrases or sentential complements as their deep structure subjects. As opposed to soal and chetil, the frequently-occurring obligatorily possessed nouns sebechel *his/her ability* and kirel *his/her obligation* can only take subject complements in deep structure. For this reason, sebechel and kirel are always followed in the surface structure by extraposed subject complements, as illustrated in the examples below:

(26) a. Ng sebechek el eko er a blim er a klukuk.
*my ability come, go your house tomorrow*
*I can come to your house tomorrow.*

b. A Droteo a sebechel el melekoi a tekoi er a Siabal.
*his ability speak Japanese*
*Droteo can speak Japanese.*

c. Ng kirem el menguiu tia el hong.
*your obligation read this book*
*You must read this book.*

d. A Toki a kirel el mesuub er a elechang.
*her obligation study now*
*Toki has to study now.*

## 3.1. Further examples of subject complements

Examination of negative sentences in PAL reveals that the negative state verb diak *isn't, doesn't exist* (past: dimlak) is a higher verb that can take either simple noun phrases or sentential complements as subjects. Note first the following negative expressions of existence:

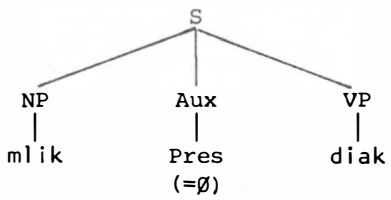
(27) a. Ng diak a mlik. *I don't have a car.*
*isn't my car*

b. Ng dimlak a ududel a Toki. *Toki didn't have any money.*
*wasn't her money*

c. Ng diak a chad er tiang. *There isn't anyone here.*
*isn't person here*

In each case, the logical subject of diak appears in sentence-final position and diak itself is preceded by the pronominal subject ng *it*. Clearly, 27a-c have been derived by the now-familiar process of subject extraposing, which leaves a pronominal trace in the original subject position. Thus, 27a has a deep structure like the following,

(28)

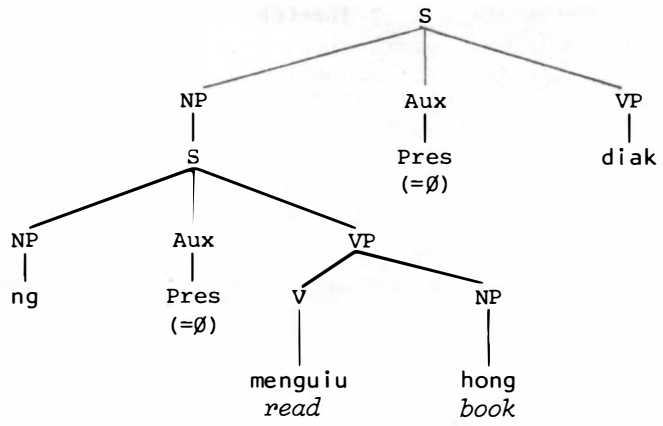



where the extraposed subject of 27a (mlik *my car*) is indeed the underlying subject of the higher (matrix) negative verb. As in the case of deep structures containing the obligatorily possessed nouns soal *his liking* and chetil *his disliking* as predicates (cf. 18-19 above), deep structures with the negative verb diak as predicate must apparently undergo subject extraposing obligatorily. Therefore, 28 must be processed or else the ungrammatical sentence 29 will result:

(29) *A mlik a diak. *I don't have a car.*

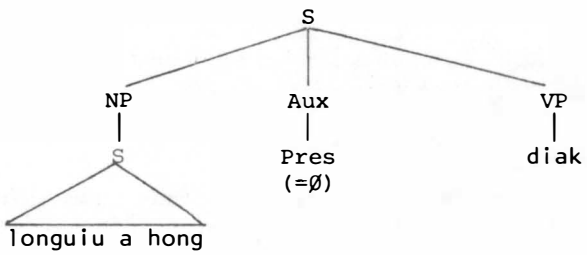
Now, if the underlying subject of diak is a whole proposition rather than a simple noun phrase, we have deep structures like the following, in which the proposition in question appears as a subject complement:

(30)




Before the subject complement of 30 is obligatorily extraposed, its subject and verb must be changed into a ***hypothetical*** verb form under the influence of the governing higher verb diak. In PAL, the hypothetical mood is required when a particular proposition is ***unrealised*** — i.e. when it is denied, supposed, assumed, imagined, wished for, etc. Because a context of denial is supplied by diak, the sequence of embedded subject ng *he* and (non-hypothetical) verb menguiu *read* must be replaced by the hypothetical verb form longuiu, which consists of a prefixed hypothetical pronoun lo- *he* and the stem of the verb menguiu.[17] Thus, 30 is transformed into

(31)




Finally, subject extraposing applies obligatorily to 31 to yield

(32) Ng diak (el) longuiu a hong. *He's not reading the books.*

In example 32, sentence-initial ng is, of course, a pronominal trace of the extraposed sentential subject. The complementiser el has been parenthesised because it is not altogether clear whether or not it really occurs in the surface structure of sentences like 32. Although all of the complement structures so far examined are clearly introduced by el in the surface structure, the phonological evidence for el in 32 is rather unconvincing. Thus, all we hear between diak and longuiu of 32 is a weak schwa ([ə]) transition. Moreover, in cases like the following,

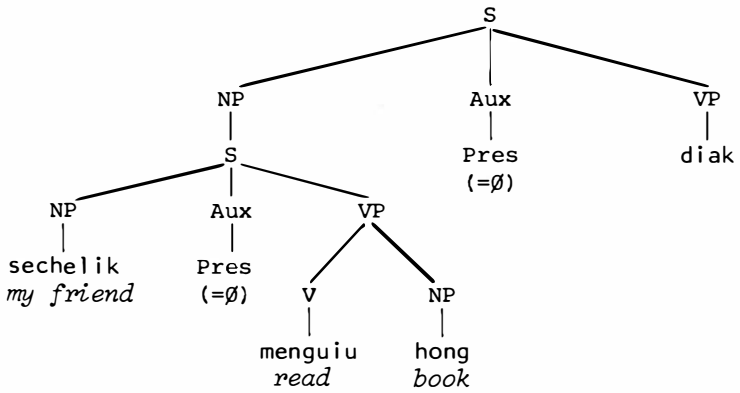
(33) a. Ng diak (el) ksecher. *I'm not sick.*
*isn't* *I-sick*

b. Ng diak (el) chobong. *You're not going.*
*isn't* *you-go*

there is no sound at all between the k of diak and the initial k or ch (glottal stop) of the immediately following hypothetical verb form. For purposes of maintaining the generalisation that all subject and object complements are introduced by el in the surface structure, it would obviously be preferable to claim that el also exists in sentences like 32 and 33a-b. But this claim would have to be made at the expense of rather ad hoc phonological rules that would either reduce el to schwa or delete it entirely. This question cannot be solved definitely at the present time and must be left open to further research.[18]

Now, suppose that the subject complement of a deep structure like 30 contains a specific third person subject rather than a pronoun such as ng *he*; in other words, consider the deep structure

(34)




Once the subject complement has been changed to the hypothetical mood, it is extraposed to yield a structure of the form

(35) ng diak (el) [sechelik longuiu hong]$_S$

The structure 35 represents an intermediate structure that must be further changed by applying either of two transformational rules. These rules operate to meet an apparent constraint that prevents a specific third person subject (e.g. sechelik *my friend* of 35) from remaining in the initial position of an extraposed subject complement. Thus, a rule of subject extraposing must apply ***within*** the extraposed complement of 35 to give

(36) Ng diak (el) longuiu a hong a sechelik.
*My friend isn't reading the books.*

Alternatively, a subject-raising rule must move the embedded subject sechelik back to sentence-initial position, where it substitutes for the pronominal trace ng. Application of subject-raising to 35 yields the following sentence:

(37) A sechelik a diak (el) longuiu a hong.
*My friend isn't reading the books.*

Some further examples parallel to 36-7 are listed below:

(38) a. Ng diak (el) loilil er a sers a rengalek.
*aren't play in garden children*
*The children aren't playing in the garden.*

b. A rengalek a diak (el) loilil er a sers.
*The children aren't playing in the garden.*

(39) a. Ng diak (el) leklou a mlid.
*isn't big our car*
*Our car isn't (that) big.*

b. A mlid a diak (el) leklou.
*Our car isn't (that) big.*

## 4. CHARACTERISTICS OF ADVERBIAL COMPLEMENTS

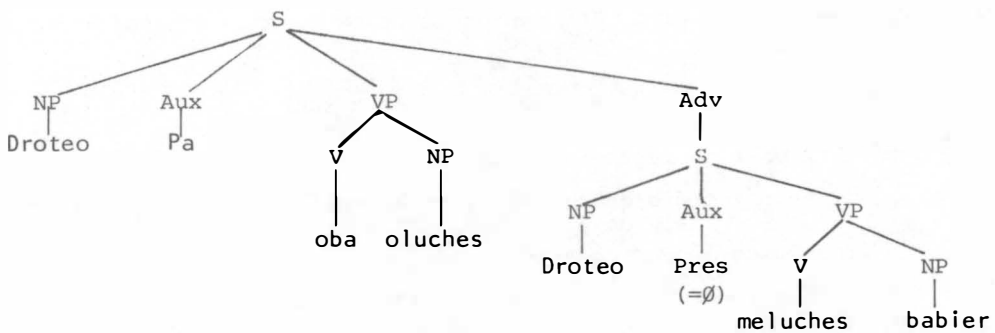
In the sections above, we have surveyed PAL subject and object complements, both of which are introduced as embedded sentences under the NP (noun phrase) node. A third major complement type to be discussed in this section does not involve embedding under NP but appears to be adverbial in nature. Observe, for example, the clauses introduced by el in the following sentences:

(40) a. A Droteo a uleba a oluches el meluches a babier.
*was using pencil write letter*
*Droteo was using a pencil to write the letter.*

b. A Droteo a milluches a babier el oba a oluches.
*was writing have, hold pencil*
*Droteo was writing the letter with (lit. holding) a pencil.*

In 40a-b, the clauses introduced by el *modify* or *specify* the preceding main clause by designating, respectively, the *purpose* of the activity of the main clause or the *instrument* with which that activity was carried out. Just as in the case of the subject and object complements discussed in 2.-3.1. above, the sequences introduced by el in 40a-b have no overtly expressed subject and contain verbs in the present tense form even though the event designated by the whole sentence took place in the past. Furthermore, the el-clauses are interpreted as having subjects which are coreferential with the subjects of the preceding clause. For these reasons, it is clear that the *purpose* and *instrument clauses* of 40a-b should be analysed in deep structure as fully specified sentential structures embedded under an adverbial node. Thus, the deep structure of 40a contains an *adverbial complement*:

(41)




In deriving 40a from 41, the familiar rule of Equi-NP Deletion must apply to the (obligatorily) identical embedded sentence subject (Droteo), and the complementiser el must be introduced in complement-initial position.

Additional examples of adverbial complements are given in the sentences below; in 42, the complements are purpose clauses, while in 43 they are instrument clauses:

(42) a. Ak ousbech a udoud el mo mechar a bail.
*I need money go buy clothes*
*I need money to (go) buy clothes.*

b. Ke okiu ker el mo er a kederang?
*you traverse where go to beach*
*What route do you take (in order) to get to the beach?*

c. Ng ngar er ngii a bilem el mo er a party?
*there are your clothes go to*
*Do you have clothes to go to the party with?*

d. A Droteo a me er a blik el me mesuub.[19]
*come to my house come study*
*Droteo is coming over to my house to study.*

(43) a. Ak milkodir a ngikel el oba a biskang.
*I killed fish have, hold spear*
*I killed the fish with a spear.*

b. A rechad er a Siabal a omengur el olab[20] a hasi.
*Japanese (people) eat have, hold chopsticks*
*The Japanese eat with chopsticks.*

## 4.1. Further examples of adverbial complements

In addition to designating purpose and instrument, PAL adverbial complements have numerous other functions, all of which involve modification or specification of the event described in the preceding main clause. The major additional types, together with any unique characteristics, are summarised below.

Any adverbial complement that specifies the conveyance used to move from one location to another is called a ***means of transportation clause***. Adverbial complements with this function always contain the existential verb ngar *exist, be (located)* followed by a ***locational phrase*** consisting of the relational word (or preposition) er *in* and a noun phrase designating any kind of conveyance — e.g. car, train, boat, etc. The sentences below contain typical means of transportation clauses:

(44) a. Ak mlo er a Siabal el ngar er a skoki.
*I went to Japan exist in plane*
*I went to Japan by plane.*

b. A Droteo a blechoel el mo er a skuul el ngar er a sidosia.
*always go to school exist in car*
*Droteo always goes to school by car.*

c. A Masaharu a mlo er a Merikel el ngar er a diall.
*went to America exist in ship*
*Masaharu went to America by ship.*

Any adverbial complement which identifies the person with whom some action is pursued is termed a ***comitative clause***. PAL comitative clauses always contain the special word obengkel *accompany, be together with*. This word, which may be related to the verb oba *have, hold, carry*, is unique in PAL in that it has the characteristics of both a verb and a noun. Thus, just like any verb prefixed with the verb marker o-,[21] obengkel has a past tense form in ule- (i.e. ulebengkel *accompanied, was together with*); furthermore, its distribution resembles that of verbs, since it directly follows the sentence subject, as in the examples below:

(45) a. Ak ulebengkel a Toki er a elii.
*I accompanied-her yesterday*
*I was with Toki yesterday.*

b. Ak ulebengkterir a resechelik er a elii.
*I accompanied-them my friends yesterday*
*I was with my friends yesterday.*

The noun-like behaviour of obengkel should also be clear from 45a-b. In these sentences, obengkel behaves like an ***obligatorily possessed noun*** in that it must always take a possessor suffix that agrees in person and number with the individual accompanying the subject in pursuit of the activity concerned. For this reason, the possessor suffix -el of ulebengkel in 45a agrees with the third person singular noun Toki, while the possessor suffix -terir of ulebengkterir in 45b agrees with the third person human plural noun resechelik *my friends*.

Now, in the sentences below, the special word obengkel occurs in adverbial complements functioning as comitative clauses:

(46) a. Ak mlo er a kedera el obengkel a Droteo.
*I went to beach accompany-him*
*I went to the beach with Droteo.*

b. Ng sebechem el mo el obengkek?
*your ability go accompany-me*
*Can you go with me?*

c. Ak mililil el obengkterir a resechelik.
*I was playing accompany-them my friends*
*I was playing with my friends.*

d. Ng soak el mo mengedub el obengkem.
*my desire go swim accompany-you*
*I want to go swimming with you.*

Because ***purpose, instrument, means of transportation,*** and ***comitative clauses*** have the common function of specifying, qualifying, or giving further information about the action or event of the main clause, they could all be described by means of a convenient cover term such as ***specifying clause***. While particular functional labels like ***purpose clause, instrument clause,*** etc. are not as readily available for the adverbial complements found in the sentences below, it is nevertheless clear that they all represent types of specifying clauses:

(47) a. A Droteo a ulureor el mengesbreber a blai.
*was working paint house*
*Droteo was working/used to work painting houses.*

b. A sechelik a mla er a diall el me er a Belau.
*my friend was in ship come to Palau*
*(lit.) My friend was in a ship coming to Palau. = My friend came to Palau by ship.*

c. A Toki a ulebengkel a Droteo el mo er a party.
*accompanied-him go to*
*(lit.) Toki accompanied Droteo going to the party. = Toki went to the party with Droteo.*

d. A Toki a mle dengchokl el kmeed er a Droteo.
*was seated near to*
*Toki was seated near Droteo.*

(48) a. Ak mirrael el mo er a Belau.
*I travelled go to Palau*
*I travelled to Palau.*

b. A Toki a rirurt el me er a blik.
*ran come to my house*
*Toki ran to my house.*

c. A katuu a riredekekl er a chedeuel a blik el mo
*cat jumped from its roof my house go*
er a kerrekar.
*into tree*
*The cat jumped from the roof of my house into the tree.*

d. A Droteo a ngilai a ilumel el me er a party.
*brought drink come to*
*Droteo brought the drinks to the party.*

e. Ak nguu a bechik el mo er a ocheraol.
*I take my wife go to money-raising party*
*I'm taking my wife to the money-raising party.*

f. Ng soak el oldurokl er a ngelekek el mo er a Merikel.
*my desire send my child go to America*
*I want to send my child to America.*

g. A sensei a ulduruklii a dempo el mo er a Saibal.
*teacher sent telegram go to Saipan*
*The teacher sent the telegram to Saipan.*

(49) a. A Droteo a lmuut el mo er a Belau er a klukuk.
*return go to Palau tomorrow*
*Droteo is going back to Palau tomorrow.*

b. A John a liluut el me er a Hawaii er a elii.
*returned come to yesterday*
*John came back to Hawaii yesterday.*

c. Ak liluut el menguiu er a hong.
*I returned read book*
*I reread the book.*

(50) a. A Toki me a Droteo a dmak el mesuub er a library.
*and together study in*
*Toki and Droteo are studying at the library.*

b. A resechelik a dilak el mengiis er a kliokl.
*my friends were together dig hole*
*My friends were digging the hole together.*

(51) a. A Toki a mla er a Merikel el ta el rak.
*was in America one year*
*Toki was in America for one year.*

b. Ak milsuub er a Siabal el eru el buil.
*studied in Japan two month*
*I studied in Japan for two months.*

(52) a. Ak mo kie er tia el mo (er a) sabadong.
*I go live here go to Saturday*
*I'll be (living) here until Saturday.*

b. Kau a cherreuek el mo er a kodall.
*you my enemy go to death*
*You're my enemy until death.*

(53) a. Ak ngilelmii a biang el rokir.
*I drank up beer all*
*I drank up the whole bottle of beer.*

b. Ke mla chemuiu a hong er kau el rokui?
*you have read book of you all*
*Have you read all your books?*

(54) a. A Droteo a milengitakl el (mle) klou a ngerel.
*was singing was big his voice*
*Droteo was singing loudly.*

b. A Toki a ulureor el kmal (mle) meringel.
*was working very was hard*
*Toki was working very hard.*

In 47a, the specifying clause el mengesbreber a blai *painting houses* narrows down or limits the scope of the activity of working expressed in the preceding main clause; in other words, the adverbial complement describes the particular kind of work involved. Similarly, in 47b-d, the specifying clauses all serve to narrow down the scope of various *states* described in the main clause. Thus, in 47d, for example, el kmeed er a Droteo *(being) near Droteo* provides specific information about where the subject (Toki) was seated.

The sentences in 48 all have main clauses containing an intransitive verb of movement (e.g. merael *travel, walk,* remurt *run,* etc.) or a transitive verb of conveyance (e.g. nguu *bring, carry,* oldurokl *send,* etc.). The accompanying adverbial complement contains a directional verb (i.e. me *come,* mo *go,* or eko *go (to the hearer's location)*) that narrows down the scope of the action designated by the preceding main clause by indicating the *direction of movement*. In a sentence like 48c, the *source* of the movement is indicated as well, in this case by the *source phrase* er a chedeuel a blik *from the roof of my house.*

In 49 and 50, main clauses containing forms of the verbs lmuut *return, do again* and dmak *be together* are followed by specifying clauses indicating, respectively, the action that is repeated or the activity that is pursued jointly by the persons designated as subjects of the sentence. As examples like 50a-b show, the subject of dmak *be together* must always be plural (e.g. Toki me a Droteo *Toki and Droteo* of 50a or resechelik *my friends* (= plural prefix re- + sechelik) of 50b).

Finally, as the groups of sentences in 51-4 illustrate, specifying clauses can function to denote a period or duration of time (51), a time limit (52), a quantity or amount (53), or the manner in which an activity is pursued (54). Many further subtypes exist, each with its own special characteristics, but unfortunately further discussion is beyond the scope of the present paper.[22]

## 5. EARLIER STUDIES ON PAL COMPLEMENTATION

Walleser (1911) presents very little on the syntactic structure of PAL complex sentences. Thus, he gives only brief recognition (133-4) to the function of el, which he explains as joining adjectives, adverbs, or clauses to (preceding or following) nouns or verbs. He provides just a few examples involving complementation, citing sentences containing specifying clauses of manner (cf. 54 above).

Many more types of complement structures are recognised by Capell (1950). Referring to the 'ligative article' el as an element that 'serves to link together words and phrases to make them practically one unit' (5-6), Capell presents (24ff.) a wide range of examples illustrating adverbial complements of the specifying variety — e.g.

(55) a. Ngii a mengitakl el ungil. (= specifying clause of manner — cf. 54)
*he sing good*
*He sings well.*

b. Ak liluut el mei. (= specifying clause following lmuut — cf. 49)
*I returned come*
*I returned/came back.*

c. Ak kie er tia el mo er a klukuk. (= specifying clause of time limit — cf. 52)
*I stay here go tomorrow*
*I'll stay here until tomorrow.*

In addition, Capell specifically mentions (32) the use of el clauses to express purpose, giving examples such as the following:

(56) a. Ng mle el omes er ngak.
*he came see me*
*He came to see me.*

b. Ak ousbech er a besebes el melechet a kekere el klalo.
*I use string tie small thing*
*I'm using string to tie up the small things.*

Finally, Capell presents examples of object clauses (following transitive verbs such as meduch *know how (to), be skilled at* and metitur *be ignorant of, not know how (to), not be capable of* — cf. 10-11 above) and extraposed subject clauses (associated with such obligatorily possessed nouns as soak *my liking,*

*what I like* and sebechek *my ability* — cf. 22a and 26a-b).

Carlson (1967), intended for Peace Corps use, includes many examples and practice drills on PAL complement structures. Broad coverage is given to extraposed subject complements (3.2-4., 4.3., and 4.5.), purpose clauses (9.5.), and comitative clauses (8.). Interesting supplementary discussions are provided on topics such as the various pronunciations of el — i.e. as [l] after words ending in vowels, as Ø after words with final l, and as [əl] otherwise (12.8.).[2 3]

Labelling el as a Verbindungsmorpheme ('linking morpheme') Pätzold (1968) gives a wide-ranging analysis of the many constructions that involve this element. Of particular relevance are Pätzold's discussions on specifying clauses of manner and direction (87ff.), purpose clauses (134-5), Nominalsätze ('nominal sentences') containing obligatorily possessed nouns of wishing, disliking, ability, etc. (i.e. sentences with extraposed subject complements) (102ff), and object complements (137).

## NOTES

1. For discussion of the word a, which introduces most PAL noun phrases, and the word er, which marks object noun phrases as specific, see Josephs 1978:2.6-7.

2. For convenience of exposition, it will be assumed in this paper that the complementiser el is transformationally introduced. Whether this analysis is correct, or whether el has a phrase structural origin and is therefore present in deep structure, will be left as an open question. Note that el is also used to introduce PAL relative clauses and to link modifying expressions to a following noun (see Josephs 1975:Chaps. 23-24).

3. The neutralisation of tense marking noted here is found in clauses accompanying various classes of implicational and non-implicational verbs (see Karttunen 1970 for further details).

4. In this sentence, the object complement following mla mo merek contains the *state verb* bechiil *married* (cf. buch *spouse*). Therefore, a literal translation of 7b would be *Droteo has finished being married.*

5. As the Eng. equivalents for the sentences of 7 indicate, the sequence mo merek + *object complement* always refers to the finishing of something on a particular instance or occasion. Thus, in 7d, for example, the speaker is asking someone on a particular occasion (e.g. before a class is to begin) whether or not he has finished smoking. Such a question is quite different in meaning from the following:

Ke mla choitii a omelamech el dekooi?
*you have give up, quit (action of) smoking cigarette*
*Have you quit/given up smoking cigarettes?*

As opposed to 7d, here the speaker is asking the person addressed whether or not he has quit or stopped the *habitual* action of smoking. The object of mla choitii *has quit* is not a complement sentence but a noun phrase containing the ***derived action noun*** omelamech *(action of) smoking* followed by the modifying expression el dekool *of cigarettes.*

6. With concrete objects, melasem can also mean *challenge* or *check (on)*, as in the following sentences:

a. Ke melasem er ngak? — *Are you challenging me?*
*you challenge me*

b. Ke mla chesemii a mlim — *Have you checked (the condition of) your car?*
*you have check your car*

7. For a detailed discussion of these verbs, see Josephs 1975:7.4.

8. The form ngilsuterir is a perfective verb (see Josephs 1975:Chapter 6.) in the past tense; the suffix -terir *them* refers to a third person human plural object and agrees, of course, with the following specific noun object resechelik *my friends.* If a specific object is omitted from a sentence like 14b, we get the following:

Ak ngilsuterir el mengetmokl er a blai.
*I helped them clean house*
*I helped them clean the house.*

Here, the first object is the pronominal suffix -terir *them* of the perfective verb form ngilsuterir *helped them*, and the second object is the object complement introduced by el.

9. The form ulderchak is a perfective verb in the past tense. In this sentence, the first object is the pronominal suffix -ak *me* of the perfective verb form. Cf. note 8 above.

10. The plural prefix re- (sometimes reduced to r-) is attached only to human nouns. See Josephs 1975:2.5 for a detailed discussion.

11. An obligatorily possessed noun is one that has no independent unsuffixed form but must always occur with a pronominal suffix indicating the possessor. Thus, we have forms like budek *my skin*, budel *his, her, its skin*, etc. but no independent word bud *skin*. Cf. Josephs 1975:3.5.

12. In the present tense there is no overt expression of the verb 'be' in copular sentences. In the past tense, however, the stative auxiliary mle *was, were*, functions as a copula. Thus, the corresponding past form of 21 would be

A Droteo a mle sensei er a ochur.
*was teacher of math*
*Droteo was a math teacher.*

13. Another class of PAL sentences in which subject extraposing is obligatory is represented by the large variety of idiomatic expressions composed of a possessed form of the abstract noun reng *heart, spirit* followed by an intransitive or stative verb (see Josephs 1975:17.4.). Such expressions, illustrated below, are used to describe emotional states, feelings, and personality traits:

a. Ng ungil a renguk. — *I'm happy.*
*good my heart*

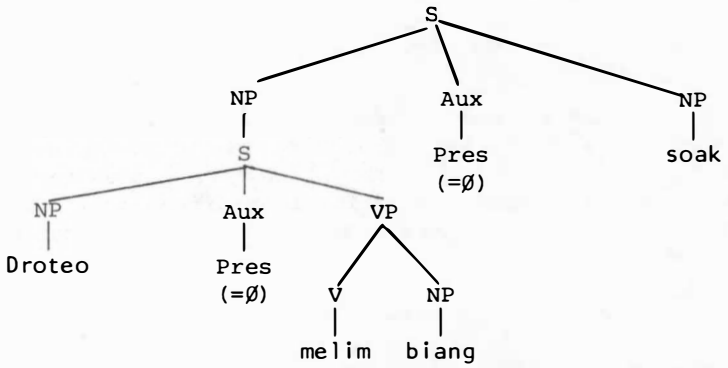
b. Ng klou a rengul a Toki. *Toki is patient.*
*big her heart*

c. Ng smecher a rengud. *We're homesick.*
*sick our* (incl.) *hearts*

d. Ng ngilasech a rengrir. *They got angry.*
*rose up their hearts*

14. Thus, according to this analysis, deep structures of the following type are ill-formed because the subject of the embedded complement (Droteo) is not identical to the possessor (*I*) of the matrix sentence possessed noun soak *my liking*.

a.




Because (a) is ill-formed, any sentences derived from it are, of course, ungrammatical:

b. *Ng soak el Droteo a melim a biang.
*(I want/would like Droteo to drink beer.)*

In order to express the intended meaning, the correct PAL sentence is

c. Ng soak a Droteo a lolim a biang.
*I want/would like Droteo to drink beer.*

Here, the sequence a lolim a biang appears to be a *conditional clause* introduced by a *if* and containing the third person singular *hypothetical* (i.e. non-real or conditional) verb form lolim *(if) he drinks* (cf. non-hypothetical melim *drink*). Thus, sentence (c), which means, literally, *I would like it if Droteo drank beer*, has a deep structure different from that of (a). For further justification of this analysis, see Josephs 1975:17.9., 19.1., and 19.4.

15. Note that ng occurs as the pronominal trace for extraposed subjects that fall into either of the following categories: (i) singular noun, human or non-human, or (ii) subject complement designating an activity, event, or proposition. By contrast, te occurs as the pronominal trace only when the extraposed subject is a plural human noun.

16. For further analysis of constructions which contrast with respect to the semantic opposition *general statement* vs. statement about a *specific occasion*, see Josephs 1975:17.8.

17. This description of hypothetical verb forms has been greatly oversimplified. For a more detailed discussion of the morphological and semantic properties of these forms, consult Josephs 1975:4.10., 4.10:1-9., and Chapter 19.

18. An alternative analysis would involve the claim that extraposed structures following diak are not introduced by the complementiser el (or any complementiser, for that matter). This would apparently be an idiosyncratic feature of diak since it is possible to find other verbs whose extraposed subject complement must contain a hypothetical verb form but is nevertheless introduced by el. Thus, Robert Gibson has pointed out (personal communication) the following sentences involving the higher verb dersta *(happen) once in a while*.

a. Ng di dersta el kbo er a chei.
*only once in a while I go fishing*
*I only go fishing once in a while.*

b. Ng dersta el lebo lsecher.
*once in a while he get sick*
*He gets sick occasionally.*

19. This sentence illustrates a pattern commonly found in sentences containing purpose clauses: if the verb of the main clause is one of the three directional verbs me *come*, mo *go*, or eko *go (to the hearer's location)*, then the following purpose clause must contain an expression of the form ***directional verb + action verb*** in which the directional verb matches that of the preceding main clause. Observe the following examples:

a. Ng kirek el eko er a blim el eko remuul a kall.
*my obligation go to your house prepare food*
*I have to come to your house to prepare the food.*

b. A ngalek a mlo er a stoa el mo omechar a kall.
*child went to store buy food*
*The child went to the store to buy food.*

For further details, see Josephs 1975:15.2.1.

20. In all instrument clauses, the verb oba *have, hold, carry* must be used in one of its (perfective) forms. Thus, in 43a oba is used with a third person singular object, while in 43b olab is required with a third person plural (non-human) object.

21. For an in-depth analysis of the PAL verb marker, see Josephs 1975:Chapter 6.

22. The amazing variety of PAL specifying clauses is set forth in greater detail in Josephs 1975:Chapter 15. In addition, direct and indirect quotation are expressed by specifying clauses introduced by el kmo *(lit.) being like (the following)* or el ua (i)se *(lit.) being like that*, as in the examples below:

a. A Droteo a dilu er ngak el kmo "Ng soak el eko er a blim.".
*said to me my desire go to your house*
*Droteo said to me, "I want to come to your house.".*

b. A Droteo a dilu er ngak el kmo ng soal el me er a blik.
*Droteo told me that he wants to come to my house.*

c. A Droteo a silebedak el kmo a demal a mla mad.
*informed me his father has die*
*Droteo informed me that his father died.*

d. Ak rirenges el kmo a Cisco a mo er a Bulabe.
*heard go to Ponape*
*I heard that Cisco is going to Ponape.*

e. Ak omdasu el kmo a Toki a mo er a Guam er a klukuk.
*I think go to tomorrow*
*I think that Toki is going to Guam tomorrow.*

f. A Toki a diak loumera el kmo a Satsko a oumlai.
*isn't believe own a car*
*Toki doesn't believe that Satsko owns a car.*

For a more complete analysis of direct and indirect quotation, see Josephs 1975:Chapter 21.

23. Carlson's decision to spell el uniformly regardless of pronunciation has been adopted in the PAL orthography used in this paper and in Josephs 1975. This orthography owes much to the following principle set forth in Carlson 1967: 12.8:

> Certain structural particles (er, el, a) have been retained in the lesson materials even though they are not always pronounced in normal speech. They are written in these lessons to help show structural relations.

## BIBLIOGRAPHY

CAPELL, Arthur

1950 *A grammar of the language of Palau.* CIMA Report 6b. (Mimeo). Washington: Pacific Science Board.

CARLSON, Clayton H.

1967 *Lessons in Palauan.* Revised edition. (Mimeo). Honolulu: University of Hawaii Peace Corps Training Center (PALI).

JOSEPHS, Lewis S.

1975 *Palauan reference grammar.* Honolulu: University Press of Hawaii.

KAJITA, Masaru

1968 *A generative-transformational study of semi-auxiliaries in present-day American English.* Tokyo: Sanseido.

KARTTUNEN, Lauri

1970 On the semantics of complement sentences. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society, April 16-18, 1970,* 328-339. Chicago Linguistic Society.

PÄTZOLD, Klaus

1968 *Die Palau-Sprache und ihre Stellung zu anderen indonesischen Sprachen.* Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 6. Berlin: Dietrich Reimer.

WALLESER, Salvator

1911 Grammatik der Palausprache. In R. Lange and A. Forke, eds *Ostasiatische Studien*, 121-231. Berlin: G. Reimer.

Josephs, L.S. "Complementation in Palauan". In Bender, B.W. editor, *Studies in Micronesian linguistics*.
**C-80:125-148**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.125*

# SCHWA IN PALAUAN

Jo-Ann Flora

The language under investigation in this study is Palauan (PAL).[1] It is spoken by approximately 13,000 people in the Palau Islands, located among the Western Caroline Islands, about 500 miles east of the island of Mindanao, Philippines. Though little work has been done along historical and comparative lines, it is generally believed that PAL should be classified as a member of the Western Austronesian division of the Austronesian language family.

Since 1900, a number of studies varying in scope and depth have been done on PAL. The Walleser grammar (1911) and dictionary (1913) are the earliest known works, and are written in a traditional style. The Capell grammar (1949) is an overall sketch of the language done in conjunction with the Coordinated Investigation of Micronesian Anthropology (CIMA). During the fifties, the McManus dictionary was compiled. It is comprehensive in nature and served as an invaluable aid in conducting the field work for this present study. Three phonemic analyses, Hsu (1960), Carlson (1968), and Flora (1969), have been done, the latter two being detailed and complete accounts of the taxonomic phonemes and their distribution throughout the language. Pätzold (1968) is a comprehensive account of the structure of PAL, but its value is questionable since it was based solely on data taken from written sources, such as Walleser (1911) and (1913) and Capell (1949). The first analysis done in the framework of generative phonology is Wilson (1972), a study of the phonology and syntax of verb affixes. Josephs (1975), a reference grammar of PAL, is an overall treatment of PAL structure written primarily for the Palauan community. Flora (1974) is a generative analysis of the major phonological processes which exist in PAL.

The particular focus of this article is the analysis of phonetic schwa in PAL, the identification of its various sources in underlying structure, and the presentation of the phonological rules needed to derive phonetic schwa from these sources. The theoretical framework in which the analysis is based is that of generative phonology as formulated by Chomsky and Halle (1968) in *The Sound Pattern of English*.

Basic to a discussion of phonetic schwa in PAL is a clear understanding of how stress is assigned to words. Below is our formulation of the stress rule:

1. Stress Assignment

$$V \rightarrow [+\text{stress}] \; / \; \underline{\quad} C_0 \langle VC_0 \rangle ]_{\langle \text{stem} \rangle} \#$$

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 149-164. *Pacific Linguistics*, C-80, 1984.

> In unsuffixed forms, stress the penultimate stem vowel. In suffixed forms, stress the rightmost vowel. When there is no penultimate stem vowel, that is, in the case of a monosyllabic stem, stress the only available vowel.

To illustrate the operation of this rule, we present below some data from the verb system of PAL. We have listed some surface forms of underlying /CVCVC/ verb stems. Although a discussion of the affixes found on these forms and the phonological rules which affect them is beyond the scope of this paper, an examination of these forms can lead to an understanding of stress placement. For each example, we give the underlying stem, the present middle, the present middle inchoative, and the future participle, in that order.[2]

| | | | |
|---|---|---|---|
| 2. | *to patch* | /tabak/ | |
| | | /m+tabak/ | mətábək |
| | | /m+tabak+a/ | mətəbəká |
| | | /tabak+l/ | təbákl |
| | *to tie* | /le?ot/ | |
| | | /m+le?ot/ | məlé?ət |
| | | /m+le?ot+a/ | mələ?ətá |
| | | /le?ot+l/ | lə?ótəl |
| | *to halve* | /doba?/ | |
| | | /m+doba?/ | mədóbə? |
| | | /m+doba?+a/ | mədəbə?á |
| | | /doba?+l/ | dəbá?əl |
| | *to shave* | /tamik/ | |
| | | /m+tamik/ | mətamík |
| | | /m+tamik+a/ | mətəmka' |
| | | /tamik+l/ | təmíkl |

These examples are representative of a vast number of underlying /CVCVC/ verb stems, and so they can be relied upon for information regarding stress assignment. Notice that the unsuffixed forms, that is, the present middle forms, are stressed on the penultimate stem vowels. Whereas the suffixed forms, that is, the present middle inchoative and future participle forms, are stressed on the rightmost vowels. In the case of the present middle inchoative, the rightmost vowel is a suffix vowel, while in the future participle, the rightmost vowel is a stem vowel.

Stress assignment operates throughout the language as it does in the verb system. Take, for example, the following nouns:

| | | | |
|---|---|---|---|
| 3. | /badu/ | bád | *rock* |
| | | bədúl | *his rock* |
| | | bədəmám | *our (excl.) rock* |
| | /keri/ | kér | *question* |
| | | kəríl | *his question* |
| | | kərəmám | *our (excl.) question* |
| | /mada/ | mád | *eyes* |
| | | mədál | *his eyes* |
| | | mədəmám | *our (excl.) eyes* |

As with verbs, these nouns are stressed on the penultimate vowel if they are unsuffixed, and on the rightmost vowel if they are suffixed.

One of the major sources of phonetic schwa is the group of underlying vowels which are subject to reduction when unstressed in certain environments. Below we give some examples from the noun system:

| | | | |
|---|---|---|---|
| 4. | /mada/ | *eyes* | mád |
| | /mada+k/ | *my eyes* | mədák |
| | /mada+mam/ | *our eyes* | mədəmám |
| | /keri/ | *question* | kér |
| | /keri+k/ | *my question* | kərík |
| | /keri+mam/ | *our question* | kərəmám |
| | /skorəs/ | *stick* | skórəs |
| | /skorəs+e+k/ | *my stick* | skərəsék |
| | /skorəs+mam/ | *our stick* | skərəsám |
| | /ʔabu/ | *ashes* | ʔáb |
| | /ʔabu+k/ | *my ashes* | ʔəbúk |
| | /ʔabu+mam/ | *our ashes* | ʔəbəmám |

Our discussion of the vowel reduction process will be concerned primarily with nouns, although the process permeates the language and affects all classes of words.

What, then, are the specifics of the vowel reduction rule? First, only stem vowels are affected. The vowels of prefixes are never reduced to schwa, even though they are unstressed. Take, for example, the following nouns which contain the instrumental prefix /o-/:

| | | |
|---|---|---|
| 5. | olámk | *razor* |
| | oləmkék | *my razor* |
| | olásəʔ | *axe* |
| | oləsəʔék | *my axe* |

Second, the applicability of the rule to stem vowels varies depending upon the quality of the vowel and its position in the stem. The unstressed stem vowels /a e o/, that is, nonhigh vowels, are regularly affected by the rule in all positions. The high vowels /i u/ are treated differently. Only those unstressed high vowels that are stem final and followed by a suffix are reduced to schwa by the vowel reduction rule. For example:[3]

| | | | |
|---|---|---|---|
| 6. | /keri+mam/ | kərəmám | *our question* |
| | /keri/ | kér | *question* |
| | /keri+l/ | kərílˀ | *his question* |
| | /badu+mam/ | bədəmám | *our rock* |
| | /badu/ | bád | *rock* |
| | /badu+l/ | bədúlˀ | *his rock* |

However, the situation is different if the unstressed high vowel is flanked on both sides by stem consonants. In this position, the high back vowel is regularly deleted. For example:

| | | | |
|---|---|---|---|
| 7. | /subad/ | súbəd | *announcement* |
| | /subad+e+k/ | spədék | *my announcement* |
| | /o+rusu/ | orús | *needle* |
| | /o+rusu+k/ | orsúk | *my needle* |

The deletion rule may be formulated as follows:

8. High Vowel Deletion

$$\begin{bmatrix} V \\ +back \\ +high \\ -stress \end{bmatrix} \rightarrow \emptyset \; / \; [X \; \underset{stem}{[+cons]} \; \_\_ \; \underset{stem}{[+cons]} \; Y]$$

When the high front vowel is flanked on both sides by stem consonants, it behaves as the most stable of vowels. Normally it is subject neither to deletion nor reduction. For example:

| | | | |
|---|---|---|---|
| 9. | /diŋa/ | díŋ | *ear* |
| | /diŋa+k/ | diŋák | *my ear* |
| | /ʔilt/ | ʔílt | *ointment* |
| | /ʔilt+e+k/ | ʔílték | *my ointment* |

There are a few examples to indicate that perhaps there is a trend toward deleting high front vowels as well as high back vowels. However there are many more examples of high front vowels which do not delete than of those that do.

In summary, then, an unstressed high back vowel flanked by two stem consonants is deleted by the High Vowel Deletion rule (8). A few high front vowels in this same environment are marked to undergo this rule exceptionally. Unstressed nonhigh stem vowels flanked by two consonants are reduced to schwa. Unstressed high vowels in stem-final position are also reduced to schwa when flanked by two consonants. The Vowel Reduction rule may be formulated as follows:[4]

10. Vowel Reduction

$$\begin{bmatrix} V \\ -stress \\ +stem \\ \langle +high \rangle \end{bmatrix} \rightarrow [-tense] \; / \; C \; \_\_ \; \langle + \rangle \; C$$

Unstressed nonhigh stem vowels flanked by two consonants are reduced to schwa. Unstressed high stem vowels flanked by two consonants also reduce to schwa if a morpheme boundary is present between the vowel and the consonant to its right, that is, if the vowel is in stem-final position.

This rule, then, can account for innumerable cases of phonetic schwa as the surface realisation of an underlying tense vowel.

We turn now to another phonological process which results in a large number of schwas in phonetic representation, that of epenthesis. Consider the following data:

| | | | |
|---|---|---|---|
| 11. | /buʔi/ | *spouse* | búʔ |
| | /buʔi+k/ | *my spouse* | bəʔík |
| | /ʔusəm/ | *beard* | ʔúsəm |
| | /ʔusəm+e+k/ | *my beard* | ʔəsəmék |

| | | |
|---|---|---|
| /ʔuri/ | *laughter* | ʔúr |
| /ʔuri+k/ | *my laughter* | ʔərík |
| /duʔa/ | *skill* | dúʔ |
| /duʔa+k/ | *my skill* | dəʔák |

It appears that in these cases the unstressed high back vowel between two stem consonants has been reduced to schwa rather than deleted. In fact, deletion has taken place, and the schwas which appear in place of the deleted vowels are the result of an epenthesis rule. Notice that in each case a glottal stop appears to the right or left of the underlying back vowel. As a rule, glottal stop does not join with any consonant to form a cluster. If through affixation or some phonological process such as vowel deletion, a glottal stop appears adjacent to a consonant, a schwa is inserted between them. For example, there is a verbal affix /-l-/ which is inserted to the right of an initial stem consonant to form the past participle of the verb. When the initial consonant is glottal stop, a schwa appears between it and the /-l-/ infix.

| | | | |
|---|---|---|---|
| 12. | /kesi/ | | *scrape* |
| | /k+l+esi/ | klés | *scraped* |
| | /balo?/ | | *shoot* |
| | /b+l+alo?/ | blálə? | *shot* |
| | /?arom/ | | *taste* |
| | /?+l+arom?/ | ?əlárəm | *tasted* |
| | /?amu/ | | *break* |
| | /?+l+amu/ | ?əlám | *broken* |

This schwa epenthesis rule may be formulated as follows:

13. Glottal Schwa Epenthesis

```
* ʔ Ø C        ə
  1 2 3  =>  1 2 3
```

The string glottal stop plus consonant and its mirror image, consonant plus glottal stop, become glottal stop plus schwa plus consonant and consonant plus schwa plus glottal stop respectively.

The derivation of a form from (11) would proceed as follows:

14.

| | bú? | bəʔík |
|---|---|---|
| | *spouse* | *my spouse* |
| | /buʔi/ | /buʔi+k/ |
| Stress | búʔi | buʔí+k |
| Final Vowel Deletion | búʔ | - |
| High Vowel Deletion | - | bʔí+k |
| Vowel Reduction | - | - |
| Glottal Schwa Epenthesis | - | bəʔí+k |
| | búʔ | bəʔík |

The process of schwa epenthesis also takes place within clusters of dental consonants, usually of the consonant-liquid type. Consider, for example, the past participle forms of dental-initial verb stems.

| | | | |
|---|---|---|---|
| 15. | /t+l+abak/ | təlábək | *patched* |
| | /s+l+ubad/ | səlúbəd | *announced* |
| | /d+l+asaʔ/ | dəlásəʔ | *carved* |

Consonant-liquid dental clusters are not permitted in final position either. For example, when the future participle marker /-l/ is suffixed to a stem ending in a consonant, we find the following forms:

| | | | |
|---|---|---|---|
| 16. | /tabak+l/ | təbákl | *to be patched* |
| | /kidib+l/ | kidíbl | *to be gathered* |
| | /leʔot+l/ | ləʔótəl | *to be tied* |
| | /tabud+l/ | təbúdəl | *to be peeled* |

Notice that when a dental consonant precedes the /-l/, a schwa intervenes. To account for this we posit the rule:

17. Dental Schwa Epenthesis

$$\emptyset \rightarrow \overset{V}{[-\text{tense}]} \; / \; \begin{bmatrix} C \\ -\text{son} \\ +\text{ant} \\ +\text{cor} \end{bmatrix} \_\_ \begin{bmatrix} C \\ -\text{son} \\ +\text{ant} \\ -\text{cor} \\ -\text{nas} \end{bmatrix}$$

There is another type of schwa epenthesis which operates in widespread fashion throughout the verb morphology in PAL. Verb stems are subject to a great deal of affixation in the formation of various tenses and aspects. We find the process of schwa epenthesis particularly productive among verb prefixes.

The first of the affixes that we will consider is the Verb Marker (VM).[5] This productive affix which appears in a vast number of verbs has the underlying form /m/, though it is realised in a variety of positions in the word and in a variety of phonetic shapes: [m], [mə], [o], and [u]. For the purposes of this article, we are concerned, obviously, with the phonetic realisation [mə] of the VM /m/, since it contains a schwa.

Consider first the following active transitive verb forms of the present middle and present imperfective aspects, in which the VM shows up as a first position prefix having the surface phonetic shape [mə-]:[6]

| 18. | | | Middle | Imperfective |
|---|---|---|---|---|
| | /dasaʔ/ | *carve* | mədásəʔ | məlásəʔ |
| | /sesob/ | *burn* | məsésəb | məlésəb |
| | /loʔad/ | *break cord* | məlóʔəd | məlóʔəd |
| | /kimud/ | *cut hair* | məkímd | məŋimd |

The VM also appears as a [mə-] prefix in some stative verbs, which resemble English adjectives. For example:

| | | | |
|---|---|---|---|
| 19. | /dakt/ | mədákt | *afraid* |
| | /duʔa/ | mədúʔ | *skilled* |
| | /ʔedi/ | məʔéd | *thirsty, shallow* |
| | /ʔuu/ | məʔúw | *shady* |
| | /rur/ | mərúr | *bashful* |

We choose to represent the VM as /m/ in underlying form, because there is a good deal of evidence to show that the schwa which appears in the [mə-] prefix is predictable and can be inserted by an epenthesis rule. For example, consider the past tense forms of the verbs in (18):

| | | | | |
|---|---|---|---|---|
| 20. | /dasaʔ/ | *carve* | mildásəʔ | millásəʔ |
| | /sesob/ | *burn* | milsésəb | millésəb |
| | /loʔad/ | *break cord* | milló?əd | milló?əd |
| | /kimud/ | *cut hair* | milkímd | milŋímd |

When the Past Tense marker /-il-/ is positioned to the right of the VM to form the past tense, the surface phonetic shape of the VM is [m-]. These facts suggest that schwa epenthesis takes place when the VM is followed by a consonant, but not when it is followed by a vowel. Our formulation of the rule will read as follows:

21. Prefix Schwa Epenthesis

$$\emptyset \rightarrow \overset{V}{[\text{-tense}]} \;/\; \# \; [\text{+cons}] + \_\_ \; [\text{+cons}]$$

Insert a schwa to the left of a consonantal segment if a single consonantal prefix precedes it.

It is easy to see how this rule would convert the forms below:

| | | | |
|---|---|---|---|
| 22. | /m+dasaʔ/ | mədásəʔ | *carve (middle)* |
| | /m+dakt/ | mədákt | *afraid* |

There are at least two other verbal prefixes which, if a schwa epenthesis rule were included in the analysis, could be represented in underlying form as single consonants, as we have suggested for the VM. A small class of stative verbs is marked on the surface by a [bə-] prefix. For example:

| | | | | |
|---|---|---|---|---|
| 23. | /ralm/ | *water* | bərálm | *watery* |
| | /sokəl/ | *ringworm* | bəsókəl | *infected with ringworm* |

Another small class of stative verbs is marked by a [kə-] prefix. For example:

| | | | |
|---|---|---|---|
| 24. | /dorom/ | kədórəm | *sharp* |
| | /debo/ | kədéb | *short* |

There is one last schwa epenthesis process we wish to discuss. It takes place in word-final position before a pause if a word ends in a consonant cluster or in two consonants separated by a schwa. When words of this shape are not followed by a pause, the schwa does not appear. Thus we have the following alternations:

| | | | |
|---|---|---|---|
| 25. | dákt | dáktə | *fear* |
| | málk | málkə | *chicken* |
| | ŋáləkˀ | ŋáləkə | *child* |
| | ʔúsəm | ʔúsəmə | *beard* |

The rule may be written as follows:

26. Phrase-final Schwa Epenthesis

$$\emptyset \rightarrow \overset{V}{[\text{-tense}]} \;/\; C(\text{ə})C\_\_\,\|$$

Since this epenthesis occurs in a very restricted environment and is completely predictable, we do not represent the schwa in the surface phonetic forms that we cite throughout this study.

Thus far, we have shown that surface phonetic schwas are present in the language as a result of two phonological processes, vowel reduction and epenthesis. It would be ideal if we could explain all occurrences of surface schwa in these ways. However, a problem arises when we consider words like the following:

| | | |
|---|---|---|
| 27. | ŋálək | *child* |
| | ŋələkék | *my child* |
| | kəmúr | *tail* |
| | kəmrík | *my tail* |
| | ʔúsəm | *beard* |
| | ʔəsəmék | *my beard* |

Each of these forms contains a schwa in the unstressed syllable of the unpossessed form. It is clear that these schwas are not epenthetic when we consider forms like the following:

| | | |
|---|---|---|
| 28. | málk | *chicken* |
| | məlkék | *my chicken* |
| | kmál | *very* |
| | mərásm | *to sew* |

These examples contain the clusters [lk#], [#km], and [sm#]. The examples in (27) have these same strings of consonants separated by a schwa: [lək#], [kəm#], and [səm#]. We must conclude that the schwas in question are not epenthetic and are not present for the purpose of breaking up unwanted consonant clusters. Therefore, they must be reductions of underlying tense vowels. However, it is not possible to determine which tense vowels underlie which schwas, since forms like those in (27) are not derivationally related to other forms in which the tense vowel appears on the surface. Since any decision about which tense vowels underlie these schwas would be an arbitrary one, we have chosen to represent them as underlying schwas.

We should at this point consider the Wilson (1972) analysis of this problem. In 2.5. of her dissertation, she describes possible ways of handling the indeterminacy of the vowels underlying these schwas. She chooses to represent such vowels as /V/, that is, a segment marked [+syllabic], an unspecified vowel. She is unconvinced that underlying forms must be fully specified and wishes to employ the principle of maximum use of phonological rules. According to this principle, her rule that reduces unstressed interconsonantal vowels would also apply to underlying /V/ and convert it to surface schwa. This well-motivated and productive rule would then be maximally employed. To posit underlying schwa would be to minimise the use of the vowel reduction rule. Another reason for Wilson's positing underlying /V/ is that she finds it unconvincing to posit underlying schwa for those vowels which are deleted rather than reduced. She is referring to certain underlying stem-final vowels which never surface, but which are required for the correct application of her stress rule. It should be pointed out that unlike the stress rule we have adopted, Wilson's rule states that stress is assigned to the penultimate vowel in a word of two or more syllables. Suffixed and unsuffixed forms are stressed in the same way in her system. Take for example the following forms:

| | | |
|---|---|---|
| 29. | bád | *rock* |
| | bədúk | *my rock* |

It is clear that the second example in (29) will not be stressed correctly according to her rule, given our underlying forms.

30. /badu/ *rock*
/badu+k/ *my rock*

However, because of her conviction that her stress rule is correct, Wilson assumes that the underlying form of the suffix in (29) is not /-k/, but /-kV/, and that after stress has been assigned to the penultimate vowel, the final unspecified vowel of the suffix is deleted by the same rule that deletes the stem-final vowel of the unsuffixed form in (29). Her derivation of these forms would proceed as follows:

31.

| | /badu/ | /badu+kV/ |
|---|---|---|
| Stress | bádu | badú+kV |
| Final Vowel Deletion | bád | badú+k |
| Vowel Reduction | - | bədú+k |
| | bád | bədúk |

Wilson treats all suffixes consisting of a single consonant in this manner, that is, by placing the unspecified vowel /V/ in word-final position, so as to ensure correct stress placement. An analysis which posits underlying schwa in this position is unattractive to her. However, there is no motiviation whatsoever for positing any vowel at all in this position, if one accepts the hypothesis that suffixed and unsuffixed forms are stressed differently, as we have demonstrated. Thus the problem of which vowel to posit in forms like the second example of (29) is completely eliminated, and we are free to focus in on the more basic problem of the indeterminacy of the vowels which underlie schwas in examples such as those in (27). Our position is the following: when a surface schwa cannot be accounted for by means of the Vowel Reduction rule (10), the Glottal Schwa Epenthesis rule (13), the Dental Schwa Epenthesis rule (17), the Prefix Schwa Epenthesis rule (26), we represent it in underlying form as a schwa. It would seem that the language is moving from a position in which the underlying quality of all schwas could be determined to a position where this is no longer true. At this point, the underlying quality of only some schwas can be determined. It is our prediction that as the language develops in the future, the number of schwas whose underlying quality can be determined will diminish, and that there will arise contrasts between schwas and tense vowels that cannot be accounted for by phonological rules. At that point, the status of underlying schwa will be more secure than at present.

There is rather solid evidence to support this prediction from the verb system, specifically from future participle forms, which are composed of the underlying verb stem plus a suffix /-l/. Below we give some examples:

32.

| | | |
|---|---|---|
| /tabak+l/ | təbákl | *patch* |
| /dakul+l/ | dəkúll | *bury* |
| /da?ob+l/ | dəŋóbl | *cover* |
| /samik+l/ | səmíkl | *peel* |
| /sesob+l/ | səsóbl | *set fire* |
| /ŗusa?+l/ | rsá?əl | *pound* |
| /ḳimud+l/ | kmúdəl | *cut hair* |
| /lo?ad+l/ | lə?ádəl | *break cord* |
| /barot+l/ | bərótəl | *hide* |

Their derivations are straightforward:

33.

| | dəkúll | bərótəl |
|---|---|---|
| | /dakul+l/ | /barot+l/ |
| Stress | dakúl+l | barót+l |
| Vowel Reduction | dəkúl+l | bərót+l |
| Dental Schwa Epenthesis | - | bərótə+l |
| Epenthesis | dəkúll | bərótəl |

In the case of stems with the underlying shape /CVCVC/, such as those in (32), it is the future participle form which alone reveals the true identity of the vowel in the last syllable. Take for example the various inflected forms of the stem /tabak/ *'to patch'*:

34.

| | |
|---|---|
| Present middle | mətábək |
| Present imperfective | məlábək |
| Present perfective | |
| singular | tobəkíy |
| plural | twábək |
| Past participle | təlábək |
| Instrumental | olábək |
| possessed | oləbəkék |
| Gerund | oməlábək |
| possessed | omələbəkék |

None of these forms gives any hint as to the underlying representation for the vowel of the last syllable of the stem. In each case the surface manifestation is schwa. The future participle [təbákl] is the only form that can give us this information. In all the other forms the vowel has been changed to schwa by the Vowel Reduction rule.

Consider now the inflected forms of those /CVCVC/ stems that have a high back vowel in the last syllable. For example:

35. /dakul/ *to bury*

| | |
|---|---|
| Present middle | mədákl |
| Present imperfective | məlákl |
| Present perfective | |
| singular | dokliy |
| plural | θmákl |
| Past participle | dəlákl |
| Instrumental | olákl |
| Possessed | oləklék |
| Gerund | oməlákl |
| Possessed | omələklék |
| Future participle | dəkúll |

If we did not have the future participle of this verb, we would have to assume that the underlying form of the stem is /dakl/. But the future participle tells us that it is /dakul/. In all the other forms the vowel /u/ has been deleted by the High Vowel Deletion rule (8).

Suppose the information contained in future participles of /CVCVC/ stems were not available to us for some reason. One consequence of this would be that a very large number of underlying forms would look very different from the way they do now. Instead of having the shape /CVCVC/ where both vowels are full vowels, they would have the shapes /CVCəC/ or /CVCC/. This would increase the number of underlying schwas in the language and strengthen the status of

this vowel as an underlying segment. It would also minimise the application of the Vowel Reduction rule and the High Vowel Deletion rule. These rules would still be needed for other forms and so would remain well motivated. However, the number of forms to which they would apply would be greatly lessened.

The state of affairs that we have just described is precisely that which we predict will exist in the near future, and for the following reasons. The future participles we have cited above in (32) are used mainly by older generations of Palauans. Rarely is a young person able to produce them in an elicitation session. Sometimes, when presented with such a form, the young person will admit that he understands it and has heard older people using it, but he does not use it himself. Among young people a new future participle suffix /-all/ has come into use. Consequently, for some verbs we find two future participles. For example:

| 36. | | | Conservative | Innovative |
|---|---|---|---|---|
| | /daŋob/ | *cover* | dəŋóbl | dəŋəbáll |
| | /ba?id/ | *break taro* | bə?ídəl | bə?ədáll |
| | /te?ib/ | *pull out* | tə?íbl | tə?əbáll |
| | /ŋerod/ | *hoist* | ŋəródəl | ŋərədáll |
| | /reŋod/ | *tie* | rəŋódəl | rəŋədáll |
| | /lo?ad/ | *break cord* | lə?ádəl | lə?ədáll |
| | /sesob/ | *set fire* | səsóbl | səsəbáll |

For some verbs we find only the innovative form, as in:

| 37. | /sikəs/ | *pole raft* | sikəsáll |
|---|---|---|---|
| | /riŋət/ | *chew* | riŋətáll |
| | /b+riid/ | *scatter* | bəridáll |

Either the conservative form never existed or it has disappeared, since in some cases both old and young people use the innovative form.

Consider what happens to forms to which the innovative suffix is added. For example:

| 38. | /daŋob/ *cover* | | |
|---|---|---|---|
| | /daŋob+all/ | daŋob+áll | dəŋəb+áll |
| | | Stress | Vowel Reduction |

Because the suffix contains a vowel, it attracts the stress which in turn causes the rightmost stem vowel to be subject to Vowel Reduction. The result of this is that no inflected form of this stem shows what vowel truly underlies the surface schwa which always appears as the rightmost vowel. Consequently, the underlying form must be /daŋəb/ rather than /daŋob/. It is only because we still find the conservative future participle [dəŋóbl] that we can with assurance posit the underlying form /daŋob/.

It is our prediction that in time all the conservative future participles will disappear, not only because they are being replaced with the innovative forms, but also because there are other syntactic constructions which can successfully convey the meaning which future participles convey. One such construction is exemplified below:

| 39. | kir+é+k | əl məlábək ər a báyl |
|---|---|---|
| | *responsibility - my* | *patch* *clothing* |
| | *I must patch the clothing.* | |

This type of sentence is used much more frequently than the one which employs the future participle, which we give below:

40. a báyl a təbákl
*clothing patch*
*The clothing should be patched.*
*The clothing ought to be patched.*

Elimination of the conservative future participles is one manifestation of a trend which had its origins in earlier stages of the development of PAL. It is easy to see the effect of this trend if we look at some words which today have the underlying forms /CVCC/ or /CVCəC/, where the schwa does not alternate with a full vowel. Through comparison with other languages, we are able to establish the fact that these words once had the shape /CVCVC/, in which both vowels were full vowels that later were deleted or reduced to schwa, depending on their quality. For example:

41.

| PAL | | | PAN |
|---|---|---|---|
| dákt | *fear* | /dakt/ | *takut |
| málk | *chicken* | /malk/ | *manuk |
| ŋálək | *child* | /ŋalək/ | *anak |
| ŋíkəl | *fish* | /ŋikəl/ | *ikan |

The present-day forms have resulted from application of the High Vowel Deletion rule and the Vowel Reduction rule to forms like those reconstructed for PAN. For example, a word like *takut lost its u by High Vowel Deletion. Perhaps there was a stage in the language when a form of this word with an u alternated with one without an u as a result of various types of affixation and stress assignment. Then it still would have been possible to represent this word as /takut/. Today this is no longer possible, since the stem meaning *fear* always appears on the surface as [dakt] or [dəkt-]. It does not have an alternate containing an [u] in the last syllable, and so must be represented as /dakt/ in a synchronic analysis.

A word like *anak had its rightmost vowel reduced to schwa by Vowel Reduction. Perhaps there was an earlier stage when this word showed up sometimes with an a and sometimes with a schwa in the last syllable. Then it would have been possible to represent this word as /ŋalak/. However, since today the stem for *child* always appears as either [ŋaləkʔ] or [ŋələk-], it must be represented as /ŋaləkʔ/ in a synchronic analysis.

In such cases we see the language moving from a stage in which some stems had the shape /CVCVC/ with two full vowels to the present stage in which they have the shapes /CVCC/ or /CVCəC/. Elimination of conservative future participles continues this trend toward reshaping underlying forms. A large number of active transitive verb stems of the shape /CVCVC/ will have to be represented as /CVCC/ or /CVCəC/ when their conservative future participles disappear, since there will no longer be a way of identifying a full vowel for the last syllable.

Thus, for stems originally of the shape /CVCVC/, we have some variety in present-day underlying forms: /CVCVC/, /CVCəC/, and /CVCC/. Gradually, as the conservative future participles disappear, there will be an increase in the number of /CVCəC/ and /CVCC/ forms, at the expense of the /CVCVC/ forms. In one sense this movement toward the reshaping of underlying forms is one of simplification. Instead of three possible underlying forms for original /CVCVC/ words, as we have today, there will be two, /CVCC/ and /CVCəC/. Furthermore,

in a disyllabic form, it will not be necessary to know which of six possible vowels belongs in the second syllable, since it will always be schwa. The Stress Assignment rule will also be simplified and will read: stress the rightmost tense vowel. There will be no need to distinguish stress placement in suffixed forms from stress placement in unsuffixed forms. Suffixed /CVCəC/ and /CVCC/ forms will receive stress on the vowel of the suffix as they do with the present rule. Recall that the only suffix without a vowel that occurs on present-day /CVCVC/, /CVCəC/, and /CVCC/ forms is the conservative future participle suffix /-l/, which will eventually disappear. An unsuffixed /CVCəC/ form will receive stress on its penultimate vowel, not because stress is penultimate in suffixed forms as in the present-day system, but because the rightmost vowel of the stem is not tense. An unsuffixed /CVCC/ form will, of course, receive stress on its only vowel.

In summary, then, we have shown that phonetic schwa in PAL has several sources. It can be the reduction of an underlying tense vowel, in which case it is derived by means of the Vowel Reduction rule. It can also be the result of any one of four epenthesis rules. In some cases, schwas appear that cannot be derived by means of the reduction rule, the epenethesis rules, or any other rule. For these cases, we have posited underlying schwas and have shown that, because of a trend that is resulting in the change of underlying forms of stems, the status of such underlying schwas is becoming stronger in the language.

## NOTES

1. The research leading to this paper was supported by grants from the Department of Linguistics of the University of California, San Diego, a Fulbright-Hays Doctoral Dissertation Research Abroad Fellowship, and two grants from the Society of the Sigma Xi. I am grateful to Margaret Langdon and Sanford Schane for their insights shared in discussions of the topics contained in this paper. I am especially indebted to the Palauan people who so generously assisted me in my study of their language and so graciously welcomed me into their community.

2. The underlying forms used in this paper are the same as those used in Flora (1974). As it is not possible to present here all the arguments for their justification, the reader is directed to Chapters Two and Three of that work.

3. The unsuffixed forms are affected by a productive rule of Final Vowel Deletion of the form:

$$[-\text{stress}] \rightarrow \emptyset \; / \; \underline{\quad} \; \#$$

A discussion of this rule may be found in Flora (1974), Chapter Two.

4. This formulation of Vowel Reduction is only the first version of the rule in Flora (1974), where a new formulation was developed (60) for the broader picture of PAL phonology presented there. The version given here is quite adequate for the purposes of this article.

5. An extensive discussion of the Verb Marker may be found in Chapter Two of Flora (1974).

6. The alternations exhibited in the stem-initial consonants of these examples are quite widespread throughout the PAL verb system. They can be accounted for by a set of rules discussed in detail in Chapter Three of Flora (1974).

## BIBLIOGRAPHY

BRANDSETTER, Renward

1916 *An introduction to Indonesian linguistics.* Translated by C.O. Blagden. London: Royal Asiatic Society.

CAPELL, Arthur

1950 *A grammar of the language of Palau.* CIMA Report 6b. (Mimeo). Washington: Pacific Science Board.

CARLSON, Clayton H.

1967 *Lessons in Palauan.* Revised edition. (Mimeo). Honolulu: University of Hawaii Peace Corps Training Center (PALI).

CHOMSKY, Noam and Morris HALLE

1968 *The sound pattern of English.* New York: Harper & Row.

DEMPWOLFF, Otto

1934-38 *Vergleichende Lautlehre des austronesischen Wörtschatzes.* 3 vols. Berlin: Reimer.

FLORA, Marie Jo-Ann

1969 Analysis of the segmental phonemes of Palauan. *Languages and Linguistics* 4:1-30.

1974 Palauan phonology and morphology. Doctoral dissertation, University of California at San Diego.

HSU, Robert

1960 Palauan phonemics. Master's thesis, Georgetown University, Washington, D.C.

JOSEPHS, Lewis S.

1975 *Palauan reference grammar.* Honolulu: University Press of Hawaii.

McMANUS, Edwin G., S.J.

1955 Word list and grammar notes - Palauan-English and English-Palauan.

PÄTZOLD, Klaus

1968 *Die Palau-Sprache und ihre Stellung zu anderen indonesischen Sprachen.* Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 6. Berlin: Dietrich Reimer.

WALLESER, Salvator

1911 Grammatik der Palausprache. In R. Lange and A. Forke, eds *Ostasiatische Studien,* 121-231. Berlin: G. Reimer.

1913 *Palau Wörterbuch, I: Palau-Deutsch, II: Deutsch-Palau, nebst einem Anhange mit einigen Sprachübungen für Anfänger.* Hong Kong: Typis Societatis Missionum ad Exteros.

WILSON, Helen Irene

1972 *The phonology and syntax of Palauan verb affixes.* (Doctoral dissertation, Department of Linguistics, University of Hawaii, Honolulu.) *Working Papers in Linguistics, University of Hawaii.*

Flora, J-A. "Schwa in Palauan". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:149-164**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.149*

# THE NOTION 'PASSIVE' IN YAPESE

John Thayer Jensen

Some years ago Schütz and Nawadra (1972) showed that the so-called 'passive verb' in Fijian (FIJ) is actually a passive participle, or as I think is an equivalent term for present purposes, a resultative adjective. When I read this paper some years after it was written I was struck by the structural resemblance to forms of verbal roots in Yapese (YAP). The resemblance is especially remarkable in light of the fact that YAP is apparently no more closely related to FIJ than it is to any other Oceanic language, and probably to any other Austronesian language.

For purposes of the present discussion, a non-defective YAP verb has three forms: transitive, intransitive/non-definite object and resultative adjective form. Some examples of these three forms are:

transitive: (1) Gu bea sey ea moor. *I am splitting bamboo.*

intransitive/non-definite object:

(2) Gu bea masay moor. *I am splitting-bamboo.*

resultative adjective:

(3) Kea say ea moor. *The bamboo is split.*

transitive: (4) Gu bea pirdeqiy ea koebreq. *I am pounding a can.*

non-definite: (5) Gu bea pardëq koebreq. *I am pounding-cans.*

resultative: (6) Kea pardëq ea koebreq. *The can is flattened.*

In the transitive sentences the verbs sey *to split something* and pirdeqiy *to pound something* are formally parallel to transitive verbs in a language such as Eng. By this I mean that the direct object of such a verb takes the normal range of modifiers, and there is no special 'voice' focus in the sentence. YAP transitive verbs are 'very transitive'. Unlike transitives in many languages, if a verb is transitive in YAP then it cannot (with very few exceptions) be also used intransitively. There are no examples like Eng. 'I am eating ham' vs. 'I am eating'. Indeed there exist YAP sentences with no overt direct object, but in these cases a definite (and anaphoric) reference is always to be understood:

(7) Gu bea sey. *I am splitting it.*

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 165-170. *Pacific Linguistics*, C-80, 1984.

In (7) the 'it' in translation is always understood in YAP and the sentence is only situationally appropriate if some referent either in the prior linguistic environment or the larger situation is available. Furthermore, this is the only way to refer to a third person direct object pronominally. There is no third singular direct object suffix as in FIJ, although in other persons and numbers a suffix exists and is obligatory.

The non-definite object sentences like (2) have been called 'incorporated object constructions' by me in discussions and in Jensen (1977). This is because syntactically the predicate of this construction consists of the pair of the verb and its object. The pair act together as an intransitive verb in YAP taking the intransitive subject number particles:

(8) Qu ra masay moor gow. *They were splitting-bamboo.*

Compare the sentence with the intransitive verb maruweel *to work*:

(9) Qu ra maruweel gow. *They were working.*

The 'dual subject number particle' gow follows the predicate, thus showing that the predicate of (8) is masay moor *splitting-bamboo*. On the other hand, the pair is shown to be equivalent to an intransitive verb by the fact that transitive verbs take the subject number particle -eew:

(10) Qu ra sey-eew ea moor. *They were splitting bamboo.*

For some verbs the non-definite object form of the verb may also be used as an ordinary intransitive verb:

(11) Gu bea maruweliy ea kaarroo. *I am working on a car.*

(12) Gu bea maruweel kaarroo. *I am doing car-work.*

(13) Gu bea maruweel. *I am working.*

It may even be the case that four forms exist in principle for YAP verbs, although this will not be assumed for purposes of the present discussion. The four forms would be: transitive, non-definite object, intransitive and resultative. One verb differentiates non-definite object and intransitive forms of the verb, though it does not contain a resultative adjective form and is thus defective:

(14) Gu bea kaay ea faloowaa. *I am eating bread.*

(15) Gu bea gum faloowaa. *I am eating-bread.*

(16) Gu bea qabiich. *I am eating.*

There may in fact be other verbs that differentiate non-definite object forms and intransitive forms of the verb but this will not be pursued here.

The resultative forms are the focus of this paper. FIJ has transitives, intransitives and passive participles. So, apparently, does YAP, although with less regularity, and in YAP the intransitive and passive participle forms are not infrequently identical, and in some cases missing entirely. Here are some examples:

| transitive | intransitive | resultative | English |
|---|---|---|---|
| sey | masay | say | *split* |
| t'aer | duum' | m'ing | *break off* |
| biliig | bil | pil | *break up* |
| giiq | mag | mag | *husk* |

| | | | |
|---|---|---|---|
| th'aeb | thuum' | maath' | *cut* |
| raeg | marag | rag | *stick into ground* |
| buqöy | maboq | maboq | *jerk loose* |
| puug | puw | puw | *rise/raise up* |
| keer | makay | kay | *dig* |
| k'afeag | mak'af | k'af | *line up* |
| luf | muul | muul | *pull loose* |
| gay | magow | paat | *find* |
| biing | maab | maab | *open* |
| l'eeg | mal'ag | l'ag | *hook* |
| yip' | wup | kaen | *shoot* |
| gabadiy | gabad | gabad | *stick* |
| gothay | magooth | magothgooth | *destroy* |
| k'adeag | mak'ad | k'ad | *join* |
| qurufeeg | moquruuf | moquruuf | *burn* |
| thikiy/thingeeg | mathig | thig | *fall over* |
| guruy | maagur | maagur | *scratch* |
| pirdeqiy | pardëq | pardëq | *pound* |
| dileqiy | madalëq | madalëq | *press* |
| liiq | maliiq | yim' | *kill/die* |
| pethuy | peeth | peeth | *join* |
| rugoey | marag | rag | *stab* |
| kuruuf | makur | kur | *pierce* |
| thoey | mathow | thow | *blow* |
| finathiy | fanaath | fanaath | *spin* |
| toelubeeg | matoeluub | toeluub | *dive* |
| lithaeg | malithaeg | liith | *dive* |
| liith | luum | n'ag | *cook* |
| l'aeng | luum'/mal'aeng | luum' | *weave* |
| dey | madaaw | madaaw | *soften with heat* |
| galik'aed | magalik'aed | malik' | *dry* |
| dareag | madaar | daar | *spread* |
| teel | mateäl | tal | *tighten* |
| faathiy | faath | faathyaath | *strip* |
| qoloey | moqoloy | moqoloy | *peel* |
| qaruy | maqär | maqärqär | *stir* |
| guchothiy | moguchoth | moguchoth | *tear* |
| guth'iy | maguth' | guth' | *strip foreskin* |
| taay | matoo | taaw | *put* |
| paag | paag | qaaw | *drop* |
| k'eeg | mak'eag | yik' | *ignite* |
| quthum | maquth | m'uuth | *sharp* |

Some forms with missing resultative are:

| transitive | intransitive | resultative | English |
|---|---|---|---|
| filaeth | filaeth | * | *spread out* |
| qudiy | maqud | * | *squeeze out* |
| doedöqöy | madoqdoq | * | *suck* |
| qiring | maroqroq | * | *steal* |
| luqag | maeluqag | * | *fight for* |
| n'aeg | man'aeg | * | *throw away* |
| qunum | maqun | * | *drink* |
| languy | malaang | * | *eat raw* |
| feek | feek | * | *carry* |

| | | | |
|---|---|---|---|
| koel | makool | * | *hold* |
| yoen' | muun | * | *throw* |
| yoet' | moed/moedoed | * | *step on* |
| girengiy | gireeng | * | *pull* |
| toey | dumow | * | *build* |
| gaed | magäd | * | *strike* |
| qan | maaq | * | *pound* |
| tunguy | matöng | * | *compose* |
| l'angagiy | l'angaag | * | *clean weeds* |
| bayagiy | bayaag | * | *clean weeds* |
| laalaang | malaalaang | * | *chase violently* |
| k'aring | mak'arang | * | *tease; cause* |
| firegiy | fireeg | * | *weave* |
| l'iing | luum'/mal'iing | * | *get water* |
| piiq | piiq | * | *push; give* |
| f'oeth | f'ooth | * | *share* |
| qayuweeg | maqayuw | | *help* |

Verbs without intransitives but with resultatives include all adjectives and therefore are very numerous:

| transitive | intransitive | resultative | English |
|---|---|---|---|
| seereeg | * | sear | *run aground* |
| weereeg | * | wear | *separate* |
| moleag | * | mool | *sleep* |
| saapeag | * | saap | *face towards* |
| chuweeg | * | chuw | *go out* |
| tuqög | * | tuuq | *start; be erect* |
| leebuguy | * | leebug | *knot* |
| buguy | * | buguy | *bend* |
| chubeeg | * | chöb | *make sound* |
| fareag | * | far | *rise* |
| chubaliy | * | chubal | *pierce* |
| langabiy | * | langab | *embrace* |
| gabaliy | * | gabael | *stick* |
| kiqeg | * | kiiq | *separate* |
| rowrow naag | * | rowrow | *red/redden* |

This list can be extended by adjectives that have a transitive with naag. (Jensen 1977:131, 221).

There are also verbs without either an intransitive or a resultative:

| transitive | intransitive | resultative | English |
|---|---|---|---|
| laeg | * | * | *cut stripes* |
| yif' | * | * | *pull from ground* |
| qid | * | * | *have sexual intercourse* |
| finey | * | * | *think* |
| qadaag | * | * | *want* |
| naang | * | * | *know* |
| tuluuf | * | * | *chase away* |

The favourite morphological pattern for a triplet of words is that exhibited by dareag/madaar/daar, in which the bare root is used for the resultative form, root with prefix ma- for intransitive, and root with suffix -eeg (in certain phonological environments -eag) or -iy for transitive. However, there are many

exceptions. This is clearly not parallel to the FIJ forms, in which transitive actives take no transitive suffix, but a third singular (or other person) object suffix, while resultatives take the suffix -i. It may be speculated that the YAP transitive suffix -iy and the FIJ passive suffix -i are cognate, but there is not in any case a uniformity within the morphology of the YAP derivational pattern comparable to that in the FIJ. I note also that Schütz and Nawadra (1972) point out in footnotes 5 and 13 that there are cases of zero suffix passives in FIJ. This is common YAP. Another question treated in Schütz and Nawadra's article is the question of whether the relationship between active and passive in FIJ is inflectional or derivational; they conclude that it is derivational. For what it is worth, the YAP relationship is clearly derivational. The fundamental parallelism that seems worth pointing out is the syntactic one, however. Both YAP and FIJ appear to have true verbs which are active, and resultative adjectives (Schütz and Nawadra call them participles, and this term is acceptable to me as well) which are passive. Do the YAP resultative forms meet Schütz and Nawadra's criteria for participles?

1. Passives act like adjectives. The following examples compare resultatives used attributively with adjectives which are not resultatives:

(17) ba naqun ni ba feal' *a good house*

(18) ba naqun ni ba magothgooth *a destroyed house*

(19) ba lëy i puw ni ba weachweach *a white stick of bamboo*

(20) ba lëy i puw ni ba say *a split stick of bamboo*

Resultatives can also be used as predicates with the stative marker ba, which is not true of transitives:

(21) Ba feal' ea rea naqun ney *This house is good*

(22) Ba magothgooth ea rea naqun ney *This house is destroyed*

(23) Ba weachweach ea rea lëy i puw ney *This stick of bamboo is white*

(24) Ba say ea rea lëy i puw ney *This stick of bamboo is split*

Schütz and Nawadra's criteria 2 through 4 are specific to FIJ and not relevant to YAP. No similar criteria are available for YAP to distinguish adjectives from ordinary intransitive verbs.

5. It is not possible in FIJ as it is in English to indicate the 'agent' of a passive form. This is equally true of YAP. There are no forms like these:

(25) *Ba say ea rea lëy i puw ney ku Tamaag. *This stick of bamboo was split by Tamaag*

It may be a legitimate question to ask why YAP resultatives are not more regularly found than FIJ ones. I do not know enough FIJ grammar to be able to say that the following suggestion is not true of FIJ. However, it is a fact that the YAP pronouns contain a member which is definite though not specific (Jensen 1977:143-144). Thus in the sentence:

(26) Yi bea sey ea puw. *They are splitting bamboo.*

the pronoun yi means something like the Eng. 'they/you non-specific'. The YAP sentence could refer to one or more people splitting bamboo, but it does not characterise a situation. It is thus definite in that a definite group of one or more people is referred to, but it is non-specific in that the number is not specified. The YAP indefinite pronoun thus serves many of the functions of a true passive. Thus (26) could be translated *Bamboo is being split.* On the other hand, of course, when a specific agent is expressed, only an active is possible. It may be that because of the existence of the YAP indefinite pronoun, pressure for preservation of the passive as a productive formation was not so great. Alternatively, it may be that FIJ made the passive formation into a regular productive pattern to compensate for lack of another mode of expression.

In any case, the main point of this paper is historical. YAP is not known to be more closely related to one Austronesian language than another. Although I have studied the language intensely for twelve and a half years, and have some acquaintance with other Austronesian languages, I have not been able to even satisfy myself that YAP is an Oceanic, as opposed to Indonesian, or, alternatively, Indonesian, as opposed to Oceanic, language (assuming those terms to be amenable of precise definition). If in fact the parallelism between YAP and FIJ described here is historical and not either fortuitous or a result of borrowing, and if indeed YAP is an isolated member of the Austronesian family tree, then the category 'passive participle' or 'resultative adjective' must have belonged to Austronesian. It may prove revealing to determine what other Austronesian languages share this trait.

## BIBLIOGRAPHY

JENSEN, John Thayer

1977 *Yapese reference grammar.* Honolulu: University Press of Hawaii.

SCHÜTZ, Albert J. and Tevita NAWARDRA

1972 A refutation of the notion "passive" in Fijian. *Oceanic Linguistics* 11/2:88-109.

Jensen, J.T. "The notion 'passive' in Yapese". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:165-170**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.165*

# PROTO-OCEANIC REFLEXES IN WOLEAIAN

Anthony F. Tawerilmang and Ho-min Sohn

## 1. GENERAL

Woleaian (WOL) is a nuclear Micronesian (MC) language spoken by some 1,400 inhabitants of the atolls of Woleai, Eauripik, Faccaulap, Elato, Lamotrek, and Ifaluk, all located in the Yap District of the Trust Territory of the Pacific Islands. The language used for comparison with Proto-Oceanic (POC) in this paper is the dialect of Woleai Atoll where Tawerilmang came from.[1] POC is the name given to the reconstructed language that comprises what have been traditionally known as the MC, Polynesian, and Melanesian groups, as over against the Indonesian or Western Austronesian. The few sets of POC reconstructions then available (which had generally not included data from MC languages) are collected with slight modifications in Grace 1969, to make a total of 698 POC lexical items.[2]

The aim of this paper is to describe the pattern of phonological evolution from POC to WOL through a comparison of the two sets of vocabulary. It is hoped that the findings presented in this paper will contribute to a larger and more significant undertaking: MC comparative linguistics. This hope is particularly strong in that WOL has been found to be a language which retains POC forms faithfully and consistently in terms of both the number of obvious cognates (see Appendix) and sound (especially vowel) reflexes.

## 2. METHODOLOGY

Since our main concern is the historical development of the phonological structure of WOL, it is imperative at the outset to draw the line between synchronic and diachronic aspects of WOL phonology. Needless to say, all rules are historical products in their origin, i.e. added in the course of time. However, addition of certain rules results in the restructuring of the underlying representations of lexical items, while addition of others has nothing to do with restructuring. Let us call the former type of rules diachronic and the latter, synchronic. For instance, the vowel a is obligatorily raised to e between two high vowels in WOL. Addition of this rule in the grammar of WOL has not yet been accompanied by any restructuring of relevant lexical items in underlying representations. Thus, the appearance of the alternation between

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 171-199. *Pacific Linguistics*, C-80, 1984.

sila- and sile- in silasE *our* (incl.) *mother* and silei *my mother* has not caused the underlying form sila *mother* to be changed. On the other hand, the change of POC *t to WOL s, as in POC *tina *mother* vs. WOL sila, has certainly resulted in restructuring of the underlying representation of relevant lexical items (e.g. *tina > sila). Therefore, POC *t > WOL s is a diachronic rule.

Suppose we directly compare POC *tina and WOL sile- *mother*, as in the form silei *my mother*, ignoring the existing synchronic alternation between a and e in WOL. We would have to describe the development as *tina > sila > sile. This description would overlook the important fact that sila and sile are automatic alternants produced by a general rule of a-raising which currently applies to all forms that meet the given environment (cf. Sohn 1972 and Bender 1973). This rule, which has no exceptions, has nothing to do with the meanings of the lexical items involved, and may be presumed to be internalised by contemporary speakers of WOL. The change *t > s, however, is neither exceptionless nor is it internalised by contemporary speakers. It may once have been a synchronic rule, but is now quite foreign to the native speaker, who does normally not even know of its existence.

Currently, there are two opposing views concerning the occurrence of restructuring. In transformational generative grammar as represented by Chomsky and Halle (1968), it seems that a rule remains synchronic as long as surface forms are derivable predictably from underlying representations even by means of powerful abstract devices. Restructuring is allowed only when no predictability is possible. In natural generative phonology, on the other hand, restructuring occurs much sooner, i.e., immediately upon the addition of a new phonological rule (e.g. Hooper 1974:121). This claim of the natural generativists is an integral part of their theoretical framework, which also includes the general abolishment of rule ordering, establishment of only one level of phonological representation, and adherence to the strong naturalness condition.[3]

Our concept of restructuring falls somewhere between these two extremes. We will not postulate any abstract devices, inasmuch as we believe that such devices are mostly far from the psychological reality of the native speaker. However, we will recognise the existence and importance of deep and surface levels of phonological representation. We will also admit a certain amount of ordering in phonological rules, not only to attain greater simplicity in the statement of the conditions associated with rules but to achieve greater generality in the statement of phonological processes. Let us take an example.

In WOL, we have a productive process called dissimilation, which, operating from right to left, raises a to e before a low vowel (Sohn 1975:31-32).[4] Thus, we have the following alternations.

(1) {[meramE] *moon*
[maremalI] *moon of*

{[temai] *my father*
[tamemamI] *our* (excl.) *father*

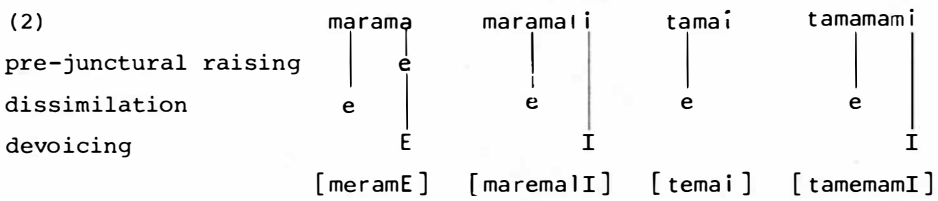
In the natural generative framework, which admits only one level of phonological representation, the underlying stem of *moon* would be either merame or marema, and that of *father* either tema or tame. However, there is no way to predict one form from the other. Both forms in each set are derivable only from a third which is never realised on the surface, i.e. marama for *moon* and tama for *father*. If, in this situation, we were to consider both surface forms in each set as lexical representations, we would miss an important phonological

generalisation (i.e. the dissimilation process) which is purely phonological and applies without exception. Moreover, to the linguistically unsophisticated native speaker of WOL, the formal difference between meramE and marema (in maremalI), for example, is not recognised, because it comes about entirely automatically, just as in allophonic alternation. We consider an allophonic variation to be a synchronic alternation. Then, there is no reason why we should not also consider purely phonological alternations such as this dissimilation process to be synchronic facts. Both phenomena are governed by the native speaker's unconscious, meaning-disregarding linguistic habits. Insofar as the native speaker of WOL recognises the two forms meramE and marema as one, it would be unreasonable to treat the dissimilation process as a historical (or diachronic) rule.

A natural corollary of considering it to be a synchronic rule is that we have to admit the existence of two levels of phonological representation. A basic premise of this paper, therefore, is that diachronic rules apply to POC forms and derive the corresponding WOL base forms, which are reconstructable from purely phonological alternations, while synchronic rules apply to WOL base forms to derive the corresponding surface forms. In other words, diachronic rules deal with the patterns of restructuring, while synchronic rules deal with the patterns of currently automatic sound alternations. Based on this premise, we draw a distinction between those processes which are partly or wholly dead and those which are completely active in purely phonological environments. If a certain phonological change were suspended prematurely or were in progress through lexical diffusion, we would consider it to be a diachronic fact, since restructuring in our sense of the term is involved in either case.

Let us go back to our examples. In order to obtain the surface forms from the reconstructed base forms marama, marama-li, tama-i, and tama-mami, we must have three general synchronic rules: prejunctural raising of a to e, dissimilation, and devoicing of the final post-consonantal simple vowel.

(2)




One significant ordering to be imposed on the above rules is that prejunctural raising must apply before dissimilation because the latter requires a surface low vowel (e.g. [a]) as its environment. On the other hand, devoicing does not have to be ordered in relation to the other two rules.

One may naturally ask, then, how we should treat WOL pairs like b (= [ɓw]) and pw, š and c, r and c, x and k, and l and n. The members of each pair are in complementary distribution in native vocabulary, in that the first occurs only as a single consonant and the second only as a geminate. Besides, when two of the first members of a given pair meet at a morpheme boundary, they are automatically replaced by the corresponding second member, as in xx → kk. For instance, notice the alternation between xaŋi *to eat it* and kkekkaŋi *to be eating it*. The only feature that differentiates x from k is [continuant]. If we take only native vocabulary into account, x and k are certainly allophonic variants. A flood of recent borrowings (especially from Japanese), however, have caused the WOL system to develop simple pw, c, k, and n which contrast with the corresponding double or fricative ones. Moreover, some neighbouring languages retain pw, c, k, and n where WOL has b, š/r, x, and l, respectively, which fact

makes the native speaker sensitive to the phonetic differences between, for example, x and k. Based on these observations, we will regard the development of b, š, r, x, and l as historical (or diachronic) facts, while the fortition of for example xx → kk is regarded as a synchronic process.

Another thorny problem concerns the treatment of semivowels. *w and *y have been reconstructed as phonemes in POC, as in *awaŋ *mouth* and *yaŋo *yellow*. In WOL, w and y are inserted before syllable-initial vowels in an entirely predictable way (see SR 1 below). In many cases, therefore, it is difficult to tell whether a given semivowel in a WOL form is inherited or has been synchronically epenthesised. However, in the case of ya:wE *mouth*, it is easy to tell that the w is an inherited one even without the evidence of POC *awaŋ, because w is never inserted between unrounded vowels. In the case of yaŋoyaŋ0 *yellow*, on the other hand, one might want to set up aŋo-aŋo as the base form, because y is inserted in the environment #___a anyway. To this word correspond two POC forms, *aŋoaŋo and *yaŋo, both meaning *yellow*, which might suggest either yaŋo-yaŋo or aŋo-aŋo as the WOL base form. Some morphophonemic alternations involving the word in question and others, however, lead us to choose yaŋo-yaŋo. Compare the pairs in (3a) and those in (3b).

(3) a.

| | |
|---|---|
| yaŋoyaŋ0 | *yellow* |
| xeyaŋoyaŋo | *make it yellow* |
| yateffasE | *beardless* |
| xeyateffesa | *shave him completely* |
| yarusa:sa | *reddish* |
| xeyarusa:sa:li | *make it bloody* |
| yaŋŋeŋŋawE | *sloppy* |
| xeyaŋŋeŋŋewa | *make him do things sloppily* |
| yaŋekuwE | *mischievous* |
| xeyaŋekuwa | *make him mischievous* |

b.

| | |
|---|---|
| yalI | *to fly* |
| xa:li | *make it fly* |
| yarE | *to get through* |
| xa:ra | *make it get through* |
| yaxiyexI | *to think* |
| xa:xiyexi | *to plan* |
| yafE | *to swim* |
| yefa:fE | *to swim vigorously* |
| yalI | *thin piece* |
| rụwa:lI | *two thin pieces* |

Notice that in (3a) y is retained after the causative prefix xa- (xe- before a by dissimilation), whereas in (3b) y does not show up after the same prefix. Further, notice in (3b) that y does not occur in the second member of a reduplicated form (e.g. yefa:fE) and after a numeral (e.g. ruwa:lI). What this means is that the y's in (3a) and those in (3b) differ in historicity. The former are already fossilised, either through inheritance from POC or through later development, while the latter are synchronically epenthesised in the given environments. The WOL base form of *yellow* should, therefore, be yaŋo-yaŋo, and not aŋo-aŋo.

A similar phenomenon is observed with w. For instance, w is epenthetic in wolɔ:l0 *to flip* and wo:r0 *fence*, in that the former may be derived from ola-ola (cf. wolati *flip it*) and the latter from oro-oro. The above observation leads us to distinguish base-form semivowels which are inherited or historically developed from surface ones which are epenthesised through synchronic processes.

## 3. SYNCHRONIC PHONOLOGICAL PROCESSES

WOL has the following phonological inventory in the native vocabulary. Notice that many sounds are limited to single or double occurrence. This is viewed as largely due to phonological diffusion through linguistic contact with neighbouring languages, and to the internal pressures to maintain structural symmetry (Sohn et al:1976).

(4) Consonants

| | | | |
|---|---|---|---|
| p | t | c (doubly) | k (doubly) |
| | | š (singly) | x (singly) |
| pw (doubly) | s | r (singly) | |
| b (singly) | | | |
| | l (singly) | | |
| m | n (doubly) | | ŋ |
| mw | | | |

Vowels

| | | |
|---|---|---|
| i | ʉ | u |
| e | əə (long) | o |
| | a | ɔɔ (long) |

Semivowels

y w

The major synchronic rules (SR's) operative in WOL are listed below, with relevant examples. Rule ordering is specified where applicable.

SR 1 (semivowel insertion)

$$\emptyset \rightarrow \begin{cases} y \;/\; \{ \#\_\_\_\bar{V}r \;\; \bar{V}r\_\_\_V \} \\ w \;/\; \{ \#\_\_\_Vr \;\; Vr\_\_\_V \} \end{cases}$$

(Vr = rounded vowel; $\bar{V}r$ = unrounded vowel)

Condition: A semivowel is not inserted before a high vowel (i, ʉ, or u) or between two identical simple vowels.

(The semivowel y is inserted before a word-initial unrounded vowel or between an unrounded vowel and any vowel. The semivowel w is inserted before a word-initial rounded vowel or between a rounded vowel and any vowel.)

(5) Examples

| | |
|---|---|
| alʊ́sʊ́ → yalʊ́sʊ́ → yalʊ́sʊ̈ (by SR 7) | *ghost* |
| ama → yama → yame (by SR 3) → ya:mE (by SR's 7 & 8) | *office* |
| afi → yafi → ya:fI (by SR's 7 & 8) | *fire* |
| ate → yate → ya:tE (by SR's 7 & 8) | *chin* |
| weriai → weriyai → weriyei (by SR 5) | *see me* |
| ia → iya → iye (by SR 3) → i:yE (by SR's 7 & 8) | *he* |
| liia → liiya → li:yE (by SR's 3 & 7) | *kill him* |
| olo → wolo- (by SR 1) | *six* |
| laloa → lalowa → lalowE (by SR's 3 & 7) | *yesterday* |
| falʊa → falʊ́wa → falʊ́wE (by SR's 3 & 7) | *island* |
| xasʊ́ʊa → xasʊ́ʊ́wa → wasʊ́:wE (by SR's 3 & 7) | *build it* |
| uaa → uwaa → uwa (by SR 7) | *fruit* |

SR 2 (a rounding)

$$a \rightarrow o \;/\; \left\{ \begin{matrix} o \\ ɔɔ \\ u \end{matrix} \right\} \; C\_\_\_\#$$

(The simple word-final vowel a is rounded to o after a back rounded vowel followed by a simple or double consonant).

(6) Examples

| | |
|---|---|
| bunna → bunno → bunnO (by SR 7) | *heart* |
| ssooŋa → ssooŋo → sso:ŋO (by SR 7) | *anger* |
| xotɔɔta → xotɔɔto → xotɔ:tO (by SR 7) | *crack* |

When, instead of a consonant, a semivowel occurs between a rounded vowel and a, SR 2 does not apply, as in paxowa → paxowE *shark* (by SR's 3 and 7).

SR 3 (prejunctural a raising: applies after SR's 1 and 2)

$$a \rightarrow e \;/\; \left\{ \begin{matrix} C \\ G \end{matrix} \right\} \; \_\_\# $$

(The simple vowel a is raised to e between a consonant or a semivowel (G = glide) and a word boundary).

(7) Examples

| | |
|---|---|
| ita# → ite → i:tE (by SR's 7 & 8) | *name* |
| afara# → yafare (by SR's 1 & 3) → yefarE (by SR's 4 & 7) | *shoulder* |
| mwara#mwarali# → mwaremwarali → mwaremwerali (by SR's 4 & 7) | *lei of* |

SR 4 (dissimilatory a raising: applies after SR 3)

$$a \rightarrow e \;/\; \_\_\_(C) \; \left\{ \begin{matrix} a \\ ɔɔ \end{matrix} \right\}$$

Condition: This rule applies from right to left.

(The simple vowel a is raised to e before a low vowel, i.e., a or ɔɔ).

(8) Examples

matamami → matemami → matemamI (by SR 7) — *our (excl.) eyes*
matai → metai — *my eyes*
xammata → xammate (by SR 3) → xemmate → xemmatE (by SR 7) — *bailer*
xatɔɔlaa → xetɔɔlaa → xetɔ:la (SR 7) — *make it bloom*
xamɔɔa → maxɔɔwa (by SR 1) → xamɔɔwe (by SR 3) → xemɔɔwe → xemɔ:wE (by SR 7) — *erase it*

SR 5 (assimilatory a raising)

a → e / Vh (C)___(C) Vh

Condition: This rule applies across word boundaries.

(The simple vowel a is raised to e between high vowels, whether or not a simple or double consonant intervenes.)

(9) Examples

maŋi#maŋi → maŋimeŋi → maŋimeŋI (SR 7) — *to think*
itai → itei — *my name*
paaʊ̆#ʂalʊ́ → paaʊ̆ʂelʊ́ → paaʊ̆ʂelʊ̄ (SR 7) — *water provisions*

An allophonic process relevant to SR's 3-5 is that e is pronounced with the lips rounded (i.e. [ə]) before or after ʊ, as illustrated in (10). However, simple [ə] is not a separate phoneme in WOL.

(10) (a) before ʊ

| surface form | pronunciation | |
|---|---|---|
| lexʊ̄ | [ləxʊ̄] | *make it tight* |
| mmweʊ́ | [mmwəʊ́] | *broken* |
| iteʊ́ | [itəʊ́] | *who?* |
| faʊ́feʊ́ | [faʊ́fəʊ́] | *to weave* |

(b) after ʊ

| surface form | pronunciation | |
|---|---|---|
| yaʊ́tE | [yaʊ́tɘ] | *current* |
| xʊ́:sE | [xʊ́:sɘ] | *octopus* |
| pa:ʊ́selʊ̄ | [pa:ʊ́səlʊ̄] | *water provisions* |

SR 6 (i assimilation)

$$i \rightarrow \left\{ \begin{array}{l} ʊ \,/\, ʊ\ (C)\_\_\_ \\ u \,/\, u\ (C)\_\_\_ \end{array} \right\} \ \#$$

(The word-final vowel i is completely assimilated to the preceding ʊ or u whether a consonant intervenes or not.)

(11) Examples

alʊ́sʊ́li → yalʊ́sʊ́li (by SR 1) → yalʊ́sʊ́lʊ́ → yalʊ́sʊ́lʊ̄ (by SR 7) → *ghost of*
xɐttui → xattuu → xattu (by SR 7) — *my finger*
maŋaaxuli → meŋaaxuli (by SR 4) → meŋaaxulu → meŋa:xulU (by SR 7) — *clothes of*

SR 7 (vowel devoicing and shortening)

$$V_{\iota} \rightarrow \left\{ \begin{array}{ll} \text{devoiced} / & \{ \begin{smallmatrix} C \\ G \end{smallmatrix} \}\_\_\#\# \\ \emptyset \quad / & V_{\iota}\_\_\#\# \end{array} \right\}$$

(A simple vowel following a consonant or a semivowel (G = glide) is devoiced before a phrase boundary; a long (geminate) vowel is shortened before a phrase boundary.)

(12) Examples

| | |
|---|---|
| imwa → imwe (by SR 3) → imwE → i:mwE (by SR 8) | *house* |
| iiaa → iiyaa (by SR 1) → i:ya | *where?* |
| iraa → ira | *branch* |

SR 8 (vowel lengthening)

$$\emptyset \rightarrow : / \# (\{\begin{smallmatrix} C \\ G \end{smallmatrix}\}) \left\{ \begin{array}{l} V\_\_ (\{\begin{smallmatrix} C \\ G \end{smallmatrix}\}) \; \underset{\circ}{V} \\ V_{\iota}\_\_V_{j} \end{array} \right\} \#$$

($\underset{\circ}{V}$ = voiceless vowel)

Condition: This rule applies only to a noun.

(In a form which consists of only two simple vowels, with one or two simple consonants or semivowels (G = glide), the first vowel is lengthened.)

(13) Examples

| | |
|---|---|
| faú → fa:ú | *stone* |
| afi → yafI (by SR's 1 & 7) → ya:fI | *fire* |
| lamwo → lamw0 (by SR 7) → la:mw0 | *lagoon* |
| ia → iya (by SR 1) → iyE (by SR's 3 & 7) → i:yE | *he* |

SR 9 (fortition)

$$\left[ \begin{array}{l} b \\ x \\ \{\begin{smallmatrix} r \\ \check{s} \end{smallmatrix}\} \\ l \end{array} \right]_{\alpha} + \alpha \rightarrow \left[ \begin{array}{l} ppw \\ kk \\ \\ cc \\ nn \end{array} \right]$$

(When doubled, the fricative b, x, r, and š are plosivised, and the flap l is nasalised.)

(14) Examples

| | |
|---|---|
| bbubbuutoxo → bbubbuutox0 (by SR 7) → ppwuppwu:tox0 | *to be coming* |
| (cf. bu:tog0 *to come*) | |
| xxaxxatapaa → xxexxatepa (by SR's 4 & 7) → kkekkatepa | *to be touching it* |
| (cf. xettapE *to touch*) | |
| rrorro → rrorr0 (by SR 7) → ccocc0 | *to decorate* |
| (cf. rosi *decorate it*) | |
| rraxo → rrax0 (by SR 7) → ccax0 | *to hug* |
| (cf. raxomi *hug it*) | |
| ššašša lúúa → šše šalúúwE (by SR's 1, 3, 4, & 7) → cceccalú:wE | |
| (cf. ša:lU *water*) | *to fill it with water* |

llůtů → llůtŮ (by SR 7) → nnůtŮ — *to be jumping*
(cf. lůtŮ *to jump*)

## 4. DIACHRONIC RULES

The POC phonemic system as reconstructable from Grace 1969 is as follows:

(15) POC Consonants:

| | | | |
|---|---|---|---|
| p | t | k | ʔ |
| mp | nt | ŋk | |
| ŋp | | | |
| | d | | |
| | nd | | |
| | s | | |
| | ns | | |
| | nj | | |
| | | R | |
| | l | r | |
| m | n | ŋ | |
| ŋm | | | |

POC Vowels:

| | | |
|---|---|---|
| | i | u |
| | e | o |
| | a | |

POC Semivowels:

| | |
|---|---|
| y | w |

The following diachronic rules (DR's) show the development from POC forms to WOL base forms. The WOL forms cited are, therefore, all base forms. Ordered rules are marked as such. Unlike synchronic rules, DR's have certain exceptions which may be regarded either as having undergone idiosyncratic changes or as later borrowings from neighbouring languages.

DR 1 (final-consonant apocope)

*C > Ø / ___#

All word-final POC consonants were dropped and no reflexes are found unless protected by a suffix of a certain kind (and hence non-final in the suffixed form).

(16) Examples

| | | |
|---|---|---|
| *p | *maʔudi(p) > maůrů | *alive* |
| | *ʔatop > aso | *thatch* |
| *t | *laŋi(t) > laŋi | *sky* |
| | *masaki(t) > mataxi | *sick, pain* |
| | *ŋkinit > xili- | *to pinch, pluck* |

| | | |
|---|---|---|
| *k | *manu(k) > malʊ | *bird, animal* |
| *ʔ | *daRa(ʔ) > ccaa | *blood* |
| | *muta(ʔ) > (m)mwuta | *to vomit* |
| *s | *manipi(s) > malifi | *thin* |
| *R | *matudu(R) > masʊrʊ | *to sleep* |
| *m | *inu(m) > ʊlʊ | *to drink* but ʊlʊmii *drink it* |
| | *ndanum > šalʊ | *fresh water* |
| | *onom > olo | *six* |
| *n | *ʔatun > asʊ | *bonito* |
| *ŋ | *awaŋ > awa | *mouth* |

Notice that ʊlʊmii *drink it* retains the final *m of *inu(m) because it is followed by an object suffix. This verb belongs to the class of so-called thematic-stem transitive verbs (Sohn 1975:125-127), in that it retains the thematic consonant -m only before a suffix (cf. ʊlʊ *to drink*).[5]

DR 2 (u centralisation)

*u > ʊ / if not preceded or followed by a POC bilabial consonant

(17) Examples

| | | |
|---|---|---|
| *ʔatun | > asʊ | *bonito* |
| *ndanu(m) | > šalʊ | *fresh water* |
| *dua | > rʊa- | *two* |
| *kku | > kkʊ | *nail, toe* |
| *kuRita | > xʊsa | *octopus* |
| *kutu | > xʊsʊ | *louse* |
| *mataku(t) | > mataxʊ | *afraid* |
| *matudu(R) | > masʊrʊ | *to sleep* |
| *natu | > laʊ | *child* |
| *panua | > falʊa | *land, island* |
| *pituʔu | > fʊsʊ | *star* |
| *RuʔaRa | > ʊa | *neck* |
| *sau(ʔ) | > taʊ- | *to pull out* |
| *susu | > tʊtʊ | *breast* |
| *ʔuna(p) | > ʊla | *fish scale, body hair* |
| *ʔuda(ŋ) | > ʊra | *lobster* |

DR 2 does not apply when *u occurs in the neighbourhood of a POC bilabial consonant, as illustrated in (18).

(18) Examples

| | | | |
|---|---|---|---|
| | *lumu | > lumwu | *seaweed, moss* |
| | *-mu | > -mwu | *your* (singular) |
| | *muta(ʔ) | > (m)mwuta | *to vomit* |
| | *namu(k) | > lamwu | *mosquito* |
| | *pua(ʔ) | > uaa | *fruit* |
| | *mpua | > bbua | *(betel)-nut* |
| D. | *puko | > uxo | *net* |
| | *puŋa | > uŋa | *ridge pole* |
| | *mputo | > buso | *navel* |
| | *putu | > utu | *tree sp. Barringtonia* |
| | *tampu | > tabu | *taboo* |
| | *tumpu(ʔ) | > subu | *to be born* |
| | *ʔumu | > umwu | *earth oven* |

As we see in *putu > utu, DR 2 does not apply to *u when it follows another u which is not subject to DR 2. Similarly, the second *u remains unchanged in *ŋapulu(?) > ŋaulu *ten.* *puti (D. *punti) > wisi *banana* might be construed as an exception to (18). However, wisi may be viewed as having been derived through usi > wisi where u is diphthongised due to the following i. A questionable exception is *tau *man, person* > tau *practitioner* where the meanings of the two words are only remotely related.

Due to the operation of DR 2, the WOL phonemic system has seen the split of POC *u into u and ʉ, thus resulting in the system of six simple vowels.

In addition to the regular shift of *u to ʉ as shown in DR 2, there are some high vowel alternations conditioned by neighbouring vowels. One such alternation is represented in DR 3, which applies fairly widely.

DR 3 (i centralisation: applies after DR 2)

*i > ʉ / ___ (C) ʉ

(19)

D *anitu > anitʉ (by DR 2) > anʉtʉ > alʉsʉ (by DR's 7 & 14) *ghost*
*iku > ikʉ (by DR 2) > ʉkʉ > ʉxʉ (by DR 9) *tail*
*inu(m) > inʉ (by DR's 1 & 2) > ʉnʉ > ʉlʉ (by DR 14) *to drink*
*liu(R) > liʉ (by DR's 1 & 2) > lʉʉ *coconut*
*pitu?u > pitʉ?ʉ (by DR 2) > pʉtʉ?ʉ > fʉsʉ (by DR's 4, 5, & 7)[6] *star*

One exception is *pitu > fisi *seven,* which may have been changed in order to avoid homophony with fʉsʉ *star.* The shift *?uluŋa > ilʉŋa *pillow* is a case of change in a direction opposite to DR 3, in that the first *u changed to i (probably by way of ʉ) in dissimilation from the second ʉ. Some other forms which show irregular high vowel reflexes include *kuli(t) > xili *skin, bark* where *u (> ʉ) changed to i before i, and *ma?udi(p) > maʉrʉ *alive* and *tuki > sʉxʉ *to pound* where *i changed to ʉ in assimilation to the preceding ʉ. Still other correspondences show mutual assimilation between a POC high vowel and a neighbouring non-high vowel, as illustrated in (20). Notice that if there is an intervening POC consonant, the assimilation occurs only when the consonant is deleted by rules to be specified later.

(20)

| | | | | |
|---|---|---|---|---|
| *au, *aou | əə | *pa?u | fəə | *to tie* |
| | | *ndau(n) | šəə | *leaf* |
| | | *paRu | (xili-)fəə | *hibiscus* |
| | | *pa?oRu | ffəə | *new* |
| *eu > ɔɔ | | *seu | tɔɔ | *rake* |
| *ai > e(e) | | *mai | me | *and, with* |
| | | *sa?it | tee-tee | *to bind* |

The above changes are not entirely regular because, for instance, we have *patu > faʉ *stone,* *sapu > taʉ *to pull out,* and *?aRus > aʉta *current,* where *au did not change to əə but followed the regular shift given in DR 2. Also, along with *seu > tɔɔ, we have *kesu > xʉʉ *back of head* where *e was completely assimilated to the following ʉ. One might be able to use some kind of rule ordering in the deletion of the consonants involved (e.g. *?, *t, *p, *R, *s) to provide an account of the vowel changes. In view of the lack of supporting data, however, such a proposal does not seem particularly attractive. In any case, the fusion of vowels added two new long vowel phonemes əə and ɔɔ to the WOL phonemic system. Remember that these long vowels do not have corresponding short counterparts.

DR 4 (glottal stop deletion)

*ʔ > Ø

The POC glottal stop *ʔ has been completely lost in all positions. Deletion of word-final *ʔ has already been accounted for by DR 1. In (21) are given examples of deletion of initial and medial *ʔ.

(21)

| | | | | |
|---|---|---|---|---|
| initial: | *ʔapaRa | > | afara | *shoulder* |
| | *ʔaro-ʔopa | > | arofa- | *love, like, miss* |
| | *ʔate | > | ase | *liver* |
| | *ʔatop | > | aso | *thatch* |
| | *ʔuda(ŋ) | > | ʉra | *lobster* |
| | *ʔumu | > | umwu | *earth oven* |
| medial: | *paʔu | > | fəə | *to tie* |
| | *daʔa(n) | > | raa | *branch* |
| | *RuʔaRu | > | ʉa | *neck* |
| | *ma-ʔanu | *to be afloat* | > maalʉ | *flood* |
| | *maʔudi(p) | > | maʉrʉ | *alive* |
| | *pituʔu | > | fʉsʉ | *star* |
| | *saʔit | > | tee-tee | *to bind* |
| | *tuʔu(d) | > | sʉʉ | *to stand up* |

DR 5 (p weakening: applies after DR 2)

$$*p > \begin{cases} Ø \;/\; \_\_\_ \;\; *u, *o \\ f \;/\; \text{elsewhere} \end{cases}$$

POC *p has been lost before a POC back vowel, whereas it has shifted to f in all other positions, as illustrated in (22).

(22) *p > Ø

| | | | | |
|---|---|---|---|---|
| | *mapo | > | mɋɔ | *to heal* |
| | *napo | > | lɔɔ | *wave, surf* |
| | *pua(ʔ) | > | uaa | *fruit* |
| D. | *puko | > | uxo | *net* |
| D. | *puŋa | > | uŋa | *ridgepole* |
| D. | *punti | > | wiši | *banana* |
| | *sapu | > | taʉ- | *to pull out* |
| | *ŋapulu(ʔ) | > | ŋaulu | *ten* |

*p > f

| | | | |
|---|---|---|---|
| *ʔapaRa | > | afara | *shoulder* |
| *api | > | afi | *fire* |
| *ʔaro-ʔopa | > | arofa- | *love, like, miss* |
| *manipi(s) | > | malifi | *thin* |
| *pada | > | faša | *pandanus* |
| *pai- | > | fa- | (reciprocal prefix) |
| *pale | > | fale | *house* |
| *panua | > | falʉa | *island, land* |
| *papine | > | faifile | *woman* |
| *patu | > | faʉ | *stone* |
| *pitu | > | fisi | *seven* |
| *tipi | > | sifi | *girdle, skirt* |

One exception to DR 5 is observed in *tupa > supa *fish poison* and *Ripa > repa *to go close*, where *p is reflected as p. In view of the abundance of reliable examples supporting DR 5, the exception may be ascribed to one of the following: (a) the POC forms may be reconstructed with *mp rather than *p; (b) they are not real correspondences; (c) the POC forms were introduced in WOL as borrowings after DR 5 was no longer active; or (d) the shift *p > f stopped prematurely before *a. If the last statement was actually the case, we would have to change DR 5 to DR 5'.

DR 5'

$$*p > \begin{cases} \emptyset & / \_\_\_ *u, *o \\ p & / \_\_\_ *a \text{ (only in certain words)} \\ f & / \text{ elsewhere} \end{cases}$$

In *pili(?) > ffili *to select* and *kapi(t) > xaffii *seize it*, *p is reflected as ff. The transitive counterpart of ffili is filii *select it*, which explains the former correspondence. The latter, however, has no explanation at present.

In *nsipo > tiwe *downward* and *tapu > tawii *conch*, deletion of *p is followed by an irregular diphthongisation of the following vowel, i.e. *o > we and *u > wii, respectively.

DR 6 (bilabial denasalisation)

$$*mp, *ŋp > \begin{cases} b & / \_\_\_ *u, *o \\ p & / \text{ elsewhere} \end{cases}$$

POC *mp and *ŋp seem to have been merged, and then denasalised in WOL to b (by way of pw) before a back (or rounded) vowel, and to p elsewhere. However, a strong case cannot be made for the two POC consonants, because there is only one correspondence available in which *ŋp occurs, as shown in (23).

(23) $\left\{ \begin{matrix} *mp \\ *ŋp \end{matrix} \right\} > b$

| | | | |
|---|---|---|---|
| | *mpo- | > boo | *smell* |
| | *mpua | > bbua | *(betel)-nut* |
| | *mpule | > bulo | *white shell, cowry* |
| | *mputo | > buso | *navel* |
| | *tampu | > tabu | *taboo* |
| | *tumpu(?) | > subu | *to be born* |
| | *ŋpoŋi | > boŋi | *night* |

*mp > p

| | | | |
|---|---|---|---|
| | *mpampa(n) | > paapa | *board, plank* |
| | *mpaya | > paa | *bait, worm* |
| D. | *tampi | > tapiya | *bowl* |

Notice in the shift *mpua > bbua that *mp is reflected as bb for reasons not statable at the moment.

DR 7 (t weakening)

$$*t > \begin{cases} t & / \_\_\_ *a \\ \emptyset & / \_\_\_ *u \text{ (only in certain words)} \\ s & / \text{ elsewhere} \end{cases}$$

POC *t is retained unchanged before the POC low vowel *a. Otherwise it has shifted to s or Ø, although predominantly s. The dropping is observed only before POC *u, and that only in a very limited number of words. One can only speculate that the words with the Ø reflex have been either influenced by or borrowed from other Trukic (TK) languages, which are regarded as having undergone a second shift of *t, i.e. t > s and s > Ø (Sohn et al 1976). This speculation is partly supported by the reflexes of POC *patu *stone*, which are fasʉ (old form) and faʉ (new form). Examples of DR 7 follow:

(24) *t > t

| | | | |
|---|---|---|---|
| | *taŋi(s) | > taŋi | *to cry* |
| | *tansi(k) | > tati | *sea* |
| | *tali | > tali | *rope* |
| | *taliŋa | > taliŋa | *ear* |
| | *tama | > tama | *father* |
| | *tano(ʔ) | > talo | *earth* |
| | *taku | > taxʉ | *back* |
| | *mata | > mata | *eye* |
| | *mataku(t) | > mataxʉ | *to be afraid* |
| | *muta(ʔ) | > (m)mwuta | *to vomit* |

*t > s

| | | | |
|---|---|---|---|
| | *tuki | > sʉxʉ | *to pound* |
| | *tumpu(ʔ) | > subu | *to be born* |
| | *tuʔu(d) | > sʉʉ | *to stand up* |
| D. | *anitu | > alʉsʉ | *ghost* |
| | *ʔatun | > asʉ | *bonito* |
| | *kutu | > xʉsʉ | *louse* |
| | *matudu(R) | > masʉrʉ | *to sleep* |
| | *pitu | > fisi | *seven* |
| | *pituʔu | > fʉsʉ | *star* |
| | *ʔate | > ase | *liver* |
| | *ʔatop | > aso | *thatch* |
| | *mate | > mase | *to die* |
| | *mputo(s) | > buso | *navel* |
| | *tika | > sixa | *bad, angry* |
| | *tina | > sila | *mother* |
| | *tipi | > sifi | *girdle, skirt* |
| | *toko(n) | > soxo | *pole, staff* |

*t > Ø

| | | | |
|---|---|---|---|
| | *motu | > mmweʉ | *to break off* |
| | *natu | > laʉ | *child* |
| | *patu | > faʉ | *stone* (old form: fasʉ) |
| | *patu | > faʉ-faʉ | *to weave* |

Exceptions to *t > t are *taʔaki *to draw water* > saaxii *extract it* and *kuRita > xʉsa *octopus*. Exceptions to *t > s are *putu > utu tree sp.: *Barringtonia* and *kato > xato *basket*.

DR 8 (dental denasalisation)

*nt, *nd > š

The two POC prenasalised dental stops *nt and *nd have merged into the palatal retroflexed fricative š in WOL, as shown in (25).

(25) *nt > š

| | | | |
|---|---|---|---|
| | *-nta | > -ša | *our* (incl.) |
| | *kinta | > -xiša | *we* (incl.) |
| D. | *punti | > wiši | *banana* |

*nd > š

| | | | |
|---|---|---|---|
| | *ndanu(m) | > šalů | *water* |
| | *ndau(n) | > šəə | *leaf* |

The reconstruction of POC *puti for *banana* is not adequate as far as the WOL (and TK) data are concerned, because *puti would have to be subject to DR 7 and would result in a wrong reflex. Therefore, *punti must be the correct reconstruction.

DR 9 (velar spirantisation)

simple *k, *ŋk > x

Simple POC *k is regularly reflected as x in WOL, while POC *kk remains unchanged. As mentioned earlier, when two x's meet at a morpheme boundary, they become kk by a synchronic process. We have only one example of *ŋk shifting to x.

(26) *k > x

| | | | |
|---|---|---|---|
| | *-ʔaki | > -axi | *cause or instrument suffix* |
| | *(dl)iki | > šixi | *small* |
| | *ika(n) | > ixa | *fish* |
| | *iku | > ůxů | *tail* |
| | *kau | > xəə | *fish hook* |
| | *kiekie | > xiexie | *pandanus* |
| | *kinta | > xiša | *we* (incl.) |
| | *ko(e) | > xo | *you* |
| | *kutu | > xůsů | *louse* |
| | *lako | > laxo | *to go* |
| | *masaki(t) | > mataxi | *sick, pain* |
| | *mataku(t) | > mataxů | *to be afraid* |
| | *puko | > uxo | *net* |
| | *toko(n) | > soxo | *staff, pole* |

*kk > kk

| | | | |
|---|---|---|---|
| | *kku | > kků | *nail, claw* |

*ŋk > x

| | | | |
|---|---|---|---|
| | *ŋkinit | > xili- | *to pinch, pluck, nip* |

One aberrant correspondence is *kali > kkeli *to dig*. The WOL form may have been derived through kakali > xaxali > xexali (dissimilatory a-raising) > xxeli > kkeli.

There are a couple of examples in which *k and *ŋk correspond to Ø. These are *suku > tůůtůů *to bathe* and *waŋka(ŋ) > waa *canoe*. If these are true cognates, DR 9 has to allow for the Ø reflex. At the moment, however, we have no strong evidence to consider them true cognates.

DR 10 (d rhotacism)

*d, *r > r

POC *d and *r have been merged to r in WOL, as illustrated in (27). Notice that there are not many examples of *r > r. In Grace (1969), *(dr) indicates that some authors reconstructed *d and others *r for the same set of correspondences.

(27) *d > r

| | | | |
|---|---|---|---|
| *-da | > | -:ra | *their* |
| *daŋa(n) | > | raa | *branch* |
| *deŋa | > | raŋa | *turmeric, yellow* |
| *doŋo | > | roŋo-roŋo | *to hear* |
| *dua | > | rʉa | *two* |
| *madama | > | marama | *moon* |
| *matudu(R) | > | masʉrʉ | *to sleep* |
| *ma?udi(p) | > | maʉrʉ | *alive* |
| *sida | > | ira | *they* |
| *?uda(ŋ) | > | ʉra | *lobster* |

*r > r

| | | | |
|---|---|---|---|
| *kari(s) | > | xeri | *to scratch, tear* |
| *raku | > | raxʉ | *to take a handful, eat clumsily* |
| *turu | > | sʉrʉ | *post* |

*(dr) > r

| | | | |
|---|---|---|---|
| *(dr)ani | > | rale | *day* |
| *(dr)odo | > | rošo | *night, darkness* |
| *si(dr)i(t) | > | siri | *semen, masturbation* |

There are exceptions to DR 10, i.e. certain occurrences of *d (not of *r) are reflected as š in WOL, as in (28).

(28)

| | | | |
|---|---|---|---|
| *(dr)odo | > | rošo | *night, darkness* |
| *pada | > | faša | *pandanus* |
| *dudu | > | ʉšʉ-ʉšʉ | *to shake, collect fruit* |
| *(dl)iki | > | šixi | *small* |

The regular source of š is *nd or *nt, as we saw in DR 8. From the WOL examples given in (28), we are tempted to modify the reconstructions in such a way that *d and *(dl) are rewritten as *nd.

DR 11 (pharyngeal weakening)

$$*R > \begin{cases} r \\ \emptyset \end{cases}$$

Without statable reasons, the POC pharyngeal *R has either merged with *d and *r to r (cf. DR 10), or been lost as shown in (29). Notice that more examples favour loss.

(29) *R > r

| | | | |
|---|---|---|---|
| D. *Ratu | > | -rasi | *1,000* |
| *Ripa | > | (xa)repa | *to go close* |
| *?apaRa | > | afara | *shoulder* |
| *tiRi *to urinate* | > | siri | *to masturbate* |

*R > Ø

| | | | |
|---|---|---|---|
| *kuRita | > | xʉsa | *octopus* |
| *ʔaRus | > | aʉta | *current* |
| *ma(R)a | > | maa | *ashamed* |
| *Ruŋmaʔ | > | imwa | *house* |
| *waRo | > | yɔɔ | *string, line* |
| *pəwoRu | > | ffəə | *new* |
| *RuʔaRa | > | ʉa | *neck* |

Here belongs the correspondence *daRa(ʔ) > ccaa *blood*. The form ccaa is assumed to have resulted from *daRa(ʔ) > ʉaRa > raʀa > rraa > ccaa (by SR 9).

DR 12 (non-high vowel fusion: applies after DR 5 & DR 11)

*ao > ɔɔ

When *a and *o met as a result of dropping of the intervening consonant, they fused with each other, resulting in the long vowel ɔɔ. Three examples are found, as in (30).

| | | | | |
|---|---|---|---|---|
| (30) | *mapo | > | mɔɔ | *to heal* |
| | *napo | > | lɔɔ | *surf, wave* |
| | *waRo | > | yɔɔ | *string, rope* |

An example parallel to (30) is *mawap > mɔɔ *to yawn*, where *awa is reflected as ɔɔ.

DR 13 (dental plosivisation)

*ns, *s > t

POC *ns and *s are both reflected as t in WOL. The two proto phonemes may have been reconstructed erroneously, as Milke has pointed out (Grace 1969). According to him, the two phonemes should be one and the same in POC.

(31) *ns > t

| | | | |
|---|---|---|---|
| *nsake | > | taxe | *upwards* |
| *nsama | > | tama | *outrigger* |
| *nsila(k) | > | tʉla | *to shine* |
| *nsiwa | > | tiwa | *nine* |
| *nsai | > | i-taʉ | *who?* |
| *ansa(n) | > | ita | *name* |
| *anse | > | ate | *chin, jaw* |
| *mansu(rR) | > | matʉ | *full (of food)* |
| *pinsa | > | fita- | *how many* |
| *pinsiko | > | fitixo | *flesh* |
| *tansi(k) | > | tati | *sea* |

*s > t

| | | | |
|---|---|---|---|
| *sili | > | tili | *to enter* |
| *sola | > | tola | *coconut blossom* |
| *susu | > | tʉtʉ | *breast* |
| *susu | > | titi | *to sew* |
| *saʔit | > | tee-tee | *to bind* |
| *sapu | > | taʉ | *to pull out* |
| *seu | > | tɔɔ | *rake* |
| *ʔasu *smoke* | > | atʉ | *appearance of smoke* |
| *ma-masa | > | mmata | *dry, low tide* |
| *masaki(t) | > | mataxi | *sick, pain* |
| *masawa | > | matawa | *strand, sea* |

There are some exceptions to DR 13, as shown in (32), where *ns and *s are reflected as s or Ø. It might be the case that the WOL words either have undergone irregular shifts due to the influence of TK languages, or are recent borrowings from them.

(32) *s > s

| | | | |
|---|---|---|---|
| *ŋase | > | ŋase | *weak* |
| *si(dr)i(t) | > | siri | *semen, masturbation* (cf. *tiRi *to spurt, urinate*) |

*ns, *s > Ø

| | | | |
|---|---|---|---|
| *nsaŋa | > | aŋa-aŋa | *a measure* |
| *nsaŋi | > | aŋi | *wind* |
| *sala(n) | > | ala | *road, path* |
| *sida | > | ira | *they* |
| *kesu | > | xʊ́ʊ́ | *back of head* |
| *tasimi | > | taimi- | *sharpen (it)* |

POC *nj, which is supposed to be Milke's nasal grade of *s, is reflected in only one convincing example. This reflex is Ø, as in *njala(n) > ala *road* (cf. *sala(n)).

DR 14 (l/n neutralisation)

simple *l, *n > l

POC *n and *l are merged as *l in WOL. The only case in which *n is retained unchanged is when it is geminate (cf. Sohn et al 1976). On the other hand, when two l's meet at a morpheme boundary, they automatically become nn (SR 9).

(33) *n > l

| | | | |
|---|---|---|---|
| *-na | > | -la | *his, her, its* |
| *namo | > | lamwo | *lagoon* |
| *namu(k) | > | lamwu | *mosquito* |
| *nana(ʔ) | > | lala | *pus* |
| *niu(R) | > | lʊ́ʊ́ | *coconut* |
| *anitu | > | alʊ́sʊ́ | *ghost* |
| *ndanu(m) | > | šalʊ́ | *fresh water* |
| *inu(m) | > | ʊ́lʊ́ | *to drink* |
| *ŋkinit | > | xili | *to pinch, pluck* |
| *ma-ʔanu *float* | > | maalʊ́ | *flood* |
| *manawa | > | malawa | *to live, breathe* |
| *manipi(s) | > | malifi | *thin* |
| *manu(k) | > | malʊ́ | *bird* |
| *panua | > | falʊ́a | *land* |
| *tano(ʔ) | > | talo | *earth* |

*l > l

| | | | |
|---|---|---|---|
| *lako | > | laxo | *go, walk* |
| *laŋi(t) | > | laŋi | *sky* |
| *laŋo | > | laŋo | *a fly* |
| *limu (*lumut) | > | lumwu | *sea-weed, moss* |
| *njala(n) | > | ala | *road* |
| *sola | > | ʈola | *coconut blossom* |
| *tali | > | tali | *rope* |
| *ŋapulu(ʔ) | > | ŋaulu | *ten* |

| | | | |
|---|---|---|---|
| *taliŋa | > | taliŋa | *ear* |
| *pale | > | fale | *house* |
| *ʔuluŋa | > | ilʉŋa | *pillow* |

DR 15 (velarisation)

$$*m > \begin{cases} mw & / \_\_\_ \ *u, *o \\ m & / \text{elsewhere} \end{cases}$$

POC *m has been split into velarised mw and plain m. Velarised mw, which occurs only before a POC round vowel, has merged with the POC *ŋm, a labiovelar.

(34) *m > mw

| | | | |
|---|---|---|---|
| *-mu | > | -mwu | *your* |
| *muʔa | > | mmwa- | *front* |
| *mudi | > | mwiri- | *behind* |
| *muta(ʔ) | > | (m)mwuta | *to vomit* |
| *ʔumu | > | umwu | *earth oven* |
| *komu | > | xumwu | *mouthful, gargle* |
| *namo | > | lamwo | *sea, lake, lagoon* |
| *limu (*lumut) | > | lumwu | *seaweed, moss* |
| *namu(k) | > | lamwu | *mosquito* |

*m > m

| | | | |
|---|---|---|---|
| *maa | > | maa | *to be ashamed* |
| *ma-ʔanu | > | maalʉ | *flood* |
| *mai | > | me | *and, with* |
| *-mami | > | -mami | *our* (excl.) |
| *manipi(s) | > | malifi | *thin* |
| *manu(k) | > | malʉ | *bird, animal* |
| *masaki(t) | > | mataxi | *sick, pain* |
| *masawa | > | matawa | *strand, sea* |
| *lima | > | lima | *five, hand* |
| *tama | > | tama | *father* |

As mentioned above, POC *ŋm appears as mw in WOL. Examples follow:

(35)

| | | | |
|---|---|---|---|
| *ŋmalo | > | mwalo | *to submerge* |
| *ŋmata | > | mwata | *worm* |
| *Ruŋma(ʔ) | > | imwa | *house* |
| *ndaŋma | > | šimwe | *head* |
| *ŋmane | > | mwaŋea- | *sibling (different sex)* |

POC *ŋ is regularly retained unchanged, as shown in (36).

(36)

| | | | |
|---|---|---|---|
| *yaŋo | > | yaŋo-yaŋo | *yellow* |
| *deŋa | > | raŋa | *turmeric, yellow* |
| *laŋi(t) | > | laŋi | *sky* |
| *paŋu | > | faŋʉ | *to be awaken* |
| *ŋpoŋi | > | boŋi | *night* |
| *taliŋa | > | taliŋa | *ear* |
| *taŋi(s) | > | taŋi | *to cry* |
| *ʔuluŋa | > | ilʉŋa | *pillow* |

The POC semivowels are retained unchanged, as in (37). Note that there is only one example of *y > y.

(37) *y > y

| | | | |
|---|---|---|---|
| | *yaŋo | > yaŋo-yaŋo | *yellow* |

*w > w

| | | | |
|---|---|---|---|
| | *madawa | > maarawa | *green, raw* |
| | *masawa | > matawa | *sea, ocean* |
| | *walu | > wali | *eight* |
| | *awaŋ | > awa | *mouth* |

One exception to (37) is *waRo > yɔɔ *string, rope* where *w has shifted to y. Another exception is *mpaya > paa *bait* where *y has been dropped between a's.

Aside from the changes effected by DR's 2, 3, and 12, the POC vowels are consistently retained unchanged, as illustrated in (38).

(38) *a > a

| | | | |
|---|---|---|---|
| | *api | > afi | *fire* |
| | *ia | > ia | *he, she, it* |
| | *panua | > falu̇a | *islands* |

*o > o

| | | | |
|---|---|---|---|
| | *kato | > xato | *basket* |
| | *onom | > olo | *six* |
| | *ŋpoŋi | > boŋi | *night* |

*i > i

| | | | |
|---|---|---|---|
| | *ika(n) | > ixa | *fish* |
| | *tipi | > sifi | *girdle, skirt* |
| | *taŋi(s) | > taŋi | *to cry* |
| | *kinta | > xiša | *we* (incl.) |

*e > e

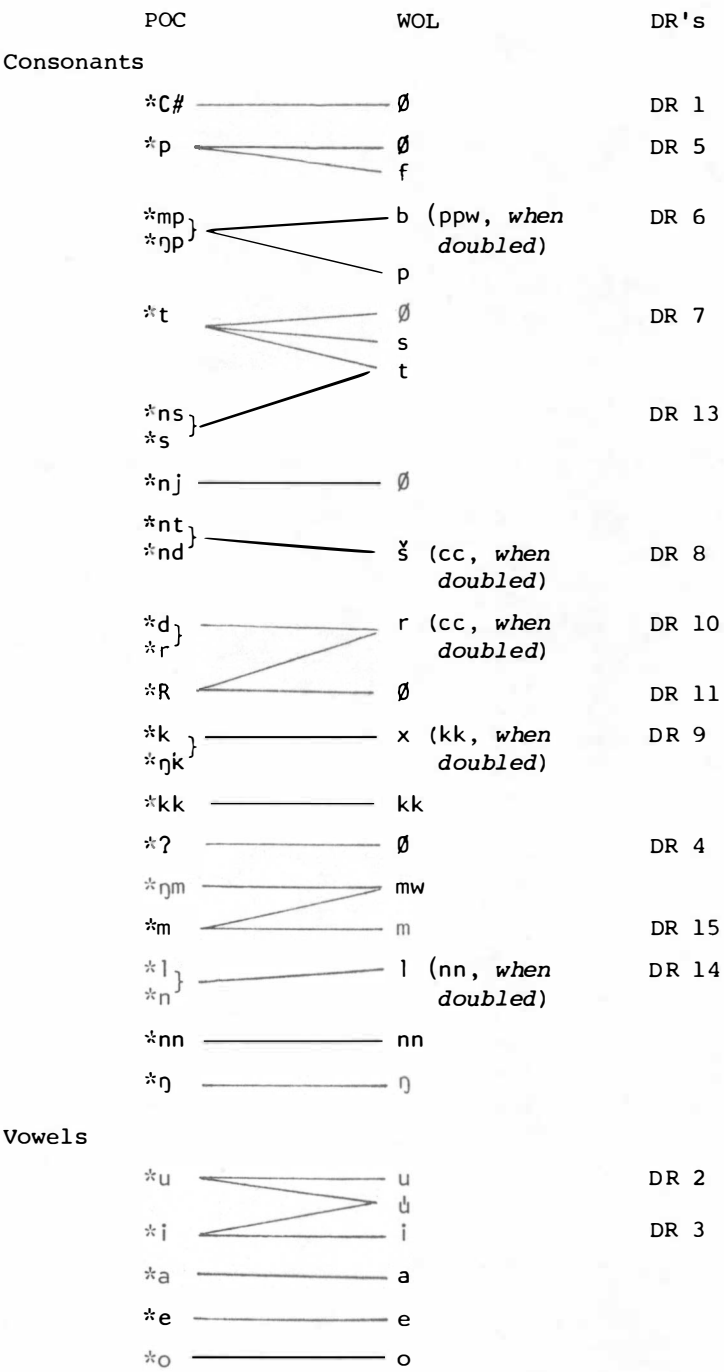
| | | | |
|---|---|---|---|
| | *anse | > ate | *chin, jaw* |
| | *kiekie | > xiexie | *pandanus* |
| | *pale | > fale | *house* |
| | *mate | > mase | *to die* |

## 5. CONCLUSION

We have tried to describe the phonological development of WOL from POC by explicitly distinguishing synchronic from diachronic aspects. Nine synchronic and fifteen diachronic rules have been presented, together with examples of lexical correspondences and some putative exceptions. As may be noted in the Appendix, there are quite a few idiosyncratic sound changes that have not been discussed. Some could be accounted for in light of the phonological environments involved, while others must await further investigation. For instance, there are a few examples of (putative) correspondences in which *n is reflected as ŋ: *nuns(io) > ŋito *squid*, *kani > xaŋi(i) *food, eat*, *ŋmane > mwaŋea- *sibling of different sex*, *nipon > ŋii *tooth*, and *ponu > oŋi *turtle*. The regular shift is *n > l in all positions, and we have no way to state a quasi-productive rule *n > ŋ, at least for the moment. We have also included a number of questionable cognates (marked with ?) in the Appendix in the hope that they might provide some important clues for future study, which should be conducted in a broader perspective, i.e. within the framework of comparative Micronesian phonology.

Let us conclude by summarising the regular sound correspondences between POC and WOL, as in (39). For the sound environments, see the fuller statements of the rules in Section 4.

(39)



*ao, etc. ——— ɔɔ DR 12

*au, etc. ——— əə

Semivowels

*y ——— y

*w ——— w

## APPENDIX. POC-WOL

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 1. | *-ʔaki | -yaxi- or -axi- | -axi | causative or instrument (affixed to verb) |
| 2. | D.*anitu | yalʊsʊ | alʊsʊ | *spirit* |
| 3. | *aŋoaŋo | yaŋoyaŋ0 | yaŋo-yaŋo | *yellow* |
| 4. | *ʔapaRa | yefarE | afara | *shoulder* |
| 5. | *api | ya:fI | afi | *fire* |
| 6. | *ʔaro-ʔopa | yarofa- | arofa- | *love, like, seek, miss* |
| 7. | *ʔaRus | yaʊtE | aʊta | *current* |
| 8. | *ansa(n) | i:tE | ita | *name* |
| 9. | *anse | ya:tE | ate | *chin, jaw* |
| 10. | D.*ʔasu | ya:tʊ *(appearance of smoke)* | atʊ | *smoke* |
| 11. | *ʔate | ya:sE | ase | *liver* |
| 12. | *ʔatop (D. *ʔato) | ya:s0 | aso | *thatch* |
| 13. | *ʔatun | ya:sʊ | asʊ | *bonito* |
| 14. | *awa | ya:wE | awa | *open space, mouth* |
| 15. | *-da | -:rE | -:ra | *their* |
| 16. | *daʔa(n) | ra | raa | *branch, twig* |
| 17. | *ndaŋma | ši:mwE *(head)* | šimwe | *head, forehead* |
| 18. | *(dr)an(i) | ra:lE | rale | *day* |
| 19. | *ndanu(m) | ša:lʊ | šalʊ | *fresh water* |
| 20. | *daRa(ʔ) | cca | ccaa | *blood* |
| 21. | *ndau(n) | šə | šəə | *leaf* |
| 22. | *deŋa | ra:ŋE | raŋa | *turmeric, yellow* |
| 23. | *(dl)iki | šixI | šixi | *small* |
| 24. | *(dr)odo | ro:š0 | rošo | *night, darkness, shadow* |
| 25. | *doŋo, ndoŋo | roŋoroŋ0 | roŋo-roŋo | *to hear* |
| 26. | D.*doŋo | ro:ŋ0 | roŋo | *rite, inheritance* |
| 27. | *dua | rʊwa- | rʊa- | *two* |
| 28. | *dudu | ʊšʊ:šʊ | ʊšʊ-ʊšʊ | *to shake, collect fruit* |
| 29. | *eno | wol0 | olo | *to lie down* |
| 30. | *ia | i:yE | ia | *he, she, it* |

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 31. | *ika(n) | i:xE | ixa | *fish* |
| 32. | *iku | ʉ:xʉ | ʉxʉ | *tail* |
| 33. | *inu(m) | ʉlʉ(m) | ʉlʉ(m) | *to drink* |
| 34. | *ka(dr)u | xerixerI | xeri-xeri | *to scratch* |
| 35. | *kai | xa:xa (Ulithi: xai) | xaa-xaa | *tree, wood* |
| 36. | *kali | kkelI | kkeli | *to dig* |
| 37. | *kami | xa:mamI | xaamami | 1st pl. excl. pronoun *(we)* |
| 38. | *kamu | xa:mi | xaamii | *ye* (2nd pl. pronoun) |
| 39. | *kani | xaŋi *(eat it)* xalE *(food)* | xaŋii, xala | *eat, food* |
| 40. | *kapi(t) | xaffi *(seize it)* | xaffii | *seize, squeeze* |
| 41. | *kari(s) | xeri *(scratch it)* | xerii | *to scratch, tear* |
| 42. | *kasup | kkutU | kkutu | *to spit* |
| 43. | *katea | xeta | xetaa | *side of canoe opposite outrigger* |
| 44. | *kato | xa:tO | xato | *basket* |
| 45. | *kau | xə | xəə | *fish hook* |
| 46. | *kau | xa:xa | xaa-xaa | *tree, stalk* |
| 47. | *kawakawa | xawexawE | xawaxawa | *fish sp. (yellow finned groper)* |
| 48. | *kawe | xə:xə | xəə-xəə | *to fasten on with straps* |
| 49. | *ke | i-xa | ixaa | *here* |
| 50. | *kesu | (xapili)xʉ | (xapili) xʉʉ | *back of head* |
| 51. | *kianto | xiyɔ | xiɔɔ | *outrigger boom* |
| 52. | *kiekie | xiyexiyE | xiexie | *pandanus* |
| 53. | *(ki)ki(t) | xʉ:xʉ | xʉʉ-xʉʉ | *to bite at, squeeze* |
| 54. | *kilala | xʉla *(know it)* | xʉlaa | *to know* |
| 55. | *ŋkinit | xilI | xili | *to pinch, pluck, nip* |
| 56. | *kinta (D. *kinta) | xi:šE | xiša | 1st incl. pronoun *(we)* |
| 57. | *ko | xo *(you!)* | xo | specifying particle, vocative particle |
| 58. | *ko(e) | xo | xo | *you, thou* |
| 59. | *komu | xumwʉ | xumwu | *mouthful, gargle* |
| 60. | *kku | kkʉ | kkʉ | *claw, nail, toe* |
| 61. | *kuli(t) | xi:lI | xili | *skin, bark* |
| 62. | *kuRita | xʉ:sE | xʉsa | *octopus* (cf. *uRita) |
| 63. | *kutu | xʉ:sʉ | xʉsʉ | *louse* |
| 64. | *lako | laxO | laxo | *to go, walk* |
| 65. | *laman | la:mwO | lamwo | *sea, lake* |
| 66. | *laŋi(t) | la:ŋI | laŋi | *sky* |
| 67. | *laŋi | ?ya:ŋI | aŋi | *wind* |
| 68. | *laŋo | la:ŋO | laŋo | *house-fly* |
| 69. | *laso | ?to:lO | tola | *genitals* |
| 70. | *lawas | lela:i | lalaai | *long* |
| 71. | *lima | li:mE | lima | *five, hand* |
| 72. | *limu | lu:mwU | lumwu | *seaweed, moss* |
| 73. | *lo | lalo, la- | lalo, la- | *in* |
| 74. | *loku | lexʉ *(make it tight)* | lexʉ | *bend, fold* |

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 75. | *lumut | lu:mwU | lumwu | *seaweed, moss* |
| 76. | *ma | me | me | *and, with* |
| 77. | *maa | ma | maa | *to be ashamed* |
| 78. | *ma-ʔanu | ma:lʊ̈ *(flood)* | maalʊ̈ | *to be afloat* |
| 79. | mada | mmašE<br>ma:rE *(preserved breadfruit)* | mmaša<br>mara | *fermented, soft, ripe* |
| 80.D. | *madama | meramE | marama | *moon* |
| 81. | *madawa | ma:rawE | maarawa | *green, raw* |
| 82. | *mai | me | me | *and, with* (cf. *ma) |
| 83. | *mala | (ni)malE *(open space)* | (ni)mala | *place* |
| 84. | *malaʔe | malemalE | mala-mala | *village, open space in village* |
| 85. | *manaŋ | ke-malE | ka-mala | *spiritual power* |
| 86. | *ŋmalo | mwalO | mwalo | *to submerge* |
| 87. | *malu | mannʊ̈ | mallʊ̈ | *soft* |
| 88. | *-mami | -mamI | -mami | *our* (excl.) |
| 89. | *manawa | melawE | malawa | *to breathe* |
| 90. | *ŋmane | mwaŋeya- | mwaŋea- | *woman's brother* |
| 91. | *manipi(s) | malifI | malifi | *thin* (cf. *mapini) |
| 92. | *manu(k) | ma:lʊ̈ | malʊ̈ | *bird, animal* |
| 93. | *mapini | malifI | malifi | *thin* (cf. *mapini(s)) |
| 94. | *mapo | mɔ | mɔɔ | *to heal (of a wound, sore)* |
| 95. | *maRa | ma | maa | *to be ashamed* (cf. *maa) |
| 96. | *ma-masa | mmatE | mmata | *dry, low tide* |
| 97. | *masaki(t) | metaxI | mataxi | *sick, pain* |
| 98. | *masawa | metawE | maʻtawa | *strand, shore, sea* (cf. *sawa) |
| 99. | *mansu(rR) | ma:tʊ̈ | matʊ̈ | *full (of food), plenty of food* |
| 100. | *mata | ma:tE | mata | *eye* |
| 101. | *ŋmata | mwa:tE | mwata | *worm* |
| 102. | *mata(ʔ) | yematE | e-mata | *raw, new* |
| 103. | *mataku(t) | metaxʊ̈ | mataxʊ̈ | *afraid* |
| 104. | *mate | ma:sE | mase | *to die, death* |
| 105. | *matudu(R) | masʊ̈rʊ̈ | masʊ̈rʊ̈ | *to sleep* |
| 106. | maʔudi(p) | maʊ̈rʊ̈ | maʊ̈rʊ̈ | *alive (of plants)* |
| 107. | *mawap | mɔ:(-li xatelʊ̈) | mɔɔ(-li xatelʊ̈) | *to yawn* |
| 108. | *moso | mottO | motta | *cooked* |
| 109. | *motu | mmweʊ̈ | mmweʊ̈ | *to break off, broken* |
| 110. | *-mu | -mwU | -mwu | *your* (sing.) |
| 111. | *muʔa | mmwa- | mmwa- | *front* |
| 112. | *mudi | mwiri- | mwiri- | *behind, after* |
| 113. | *muta(ʔ) | (m)mwutO | (m)mwuta | *to spit, vomit* |
| 114. | *muntu | mwošomwo:šO | mwošo-mwoošo | *severed, cut short* (cf. *motu) |
| 115. | *-na | -lE | -la | *his, her, its* |
| 116. | *na | lE (immed. fut.) | le | *sign of future tense* |
| 117. | *namo | la:mwO | lamwo | *lagoon* |
| 118. | *namu(k) | la:mwU | lamwu | *mosquito* |

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 119. | *nana(?) | la:lE | lala | *pus* |
| 120. | *napo | lɔ | lɔɔ | *surf, wave* |
| 121. | *natu | la:ʉ | laʉ | *child, offspring* |
| 122. | *nipi(s) | malifilifI | malifilifi | *thin* (cf. *manipi(s)) |
| 123. | *nipon | ŋi | ŋii | *tooth* |
| 124. | *niu(R) | lʉ | lʉʉ | *coconut* |
| 125. | *noRa | lalowE | laloa | *yesterday* |
| 126. | *nua | la-, le- | la-, le- | *inside* |
| 127. | *nuns(io) | ŋi:tO | ŋito | *squid* |
| 128. | *ŋapulu(?) | ŋaulU | ŋaulu | *ten* |
| 129. | *ŋase | ?ŋasE | ŋase | *weak, exhausted, lame* |
| 130. | *onom (D. ono) | wo:lO | olo | *six* |
| 131. | *pada | fa:šE | faša | *pandanus* |
| 132. | *pai- | fa-(fe-) | fa- | reciprocal prefix |
| 133. | *paka- | xa- | xa- | causative prefix |
| 134. | *pakiwak | paxowE | paxowa | *shark* |
| 135. | *pale | falE | fale | *house* |
| 136. | *palisi | fatilI | fatili | *grass* |
| 137. | *panua | falʉwE | falʉa | *land, earth, village* |
| 138. | *paŋ(ou)(n), (D. *paŋu) | faŋʊ̈ | faŋʉ | *awaken, arouse* |
| 139. | *pa?oRu | ffə | ffəə | *new* |
| 140. | *papa | fa:- | faa- | *under, downwards* |
| 141. | *mpampa(n) | pa:pE | paapa | *board, plank, flat* |
| 142. | *papine | faifilE (archaic) | faifile | *woman* |
| 143. | *paRa | yefarE | afara | *shoulder* (cf. *?apaRa) |
| 144. | *paRi | faiyE | faia | *stingray* |
| 145. | *paRu | xili-fə | xili-fəə | *hibiscus* |
| 146. | *pat | fa:- | faa- | *four* |
| 147. | *patu | fa:ʉ | faʉ | *stone* |
| 148. | *patu | faʉfeʉ | faʉ-faʉ | *to weave* |
| 149. | *pa?u | fə:- | fəə- | *to tie, bind* |
| 150. | *mpaya | pa | paa | *bait, worm* |
| 151. | *pe- | fa *(which, where)* | faa | *where?* |
| 152. | *pi(dr)i | ffišI*(snap)* | ffiši | *to fold, twist, sprain* |
| 153. | *pili(?) | ffilI | ffili | *to select, choose, pick up* |
| 154. | pine | faifilE (archaic) | faifile | *woman* |
| 155. | *pinsa | fita- | fita- | *how many* |
| 156. | *pinsiko | fitixO | fitixo | *flesh* |
| 157. | *pitu | fi:sI | fisi | *seven* |
| 158. | *pitu?u (D. *pitu?o) | fʉ:sʊ̈ | fʉsʉ | *star* |
| 159. | D.*mpo- | bo | boo | *smell* |
| 160. | *ponu | wo:ŋI | oŋi | *turtle* |
| 161. | *ŋpoŋi | bo:ŋI | boŋi | *night* |
| 162. | *ponse | ?fatʉlE | fatʉla | *paddle* |
| 163. | *pua(?) | uwa | uaa | *fruit* |
| 164. | *mpua | bbuwE | bbua | *(betel) nut* |
| 165. | *puki | wexI *(turn)* | wexi | *return* |
| 166. | D.*puko | u:xO, | uxo | *net* |
| 167. | *mpule | bulO, ubilI | bulo, ubili | *white shell, cowry* |

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 168. | *mpulu(t) | bilisE | bilisa | *gum, sap, glue* |
| 169. | D.*puŋa | u:ŋ0 | uŋa | *ridgepole* |
| 170. | *mpu(dr)i(t) | bʊ́rʊ́wE | bʊ́rʊ́a | *buttocks, excrement* |
| 171. | *mpuso(s) | bu:s0 | buso | *navel* (cf. *mputo) |
| 172. | *puti (D. *punti) | wi:šI | wiši | *banana* |
| 173. | *mputo | bu:s0 | buso | *navel, anus* (cf. *mpuso(s)) |
| 174. | *putu | u:tU | utu | *tree sp.: Barringtonia* |
| 175. | *raku | raxʊ̄ | raxʊ́ | *take a handful, eat clumsily* |
| 176. | *Ratu | (saŋe)rasI | (saŋe)rasi | *one thousand* |
| 177. | *Ripa | repE | repa | *to go close* |
| 178. | *Ru?a | ʊ́:wE | ʊ́a | *neck* |
| 179. | Ruŋma(?) | i:mwE | imwa | *house* |
| 180. | *nsanda | ?texašE | taxaša | *to rise (sun, moon), to ascend* |
| 181. | *nsai | itʊ́ | i-tʊ́ | *who?* |
| 182. | *sa?it | te:te | tee-tee | *to bind, restrain* |
| 183. | *nsake | taxE | taxe | *upwards, to climb* |
| 184. | *sala(n) | ya:lE | ala | *road, path* |
| 185. | *njala(n) | ya:lE | ala | *road, path* |
| 186. | *nsama | ta:mE | tama | *outrigger* |
| 187. | *nsaŋa | ?yeŋa:ŋE | aŋa-aŋa | *a measure, a span* |
| 188. | *nsaŋi | ?ya:ŋI | aŋi | *wind* |
| 189. | *sapu | taʊ́- | taʊ́- | *to pull out, take to pieces* |
| 190. | *sau(?) | tta:wE | ttaawa | *outside, far off* |
| 191. | *-sawa | ta:wE *(channel)* | tawa | *strand, shore, sea* (cf. *masawa) |
| 192. | *seu | tɔ | tɔɔ | *to rake, sweep, scratch* |
| 193. | *sida | i:rE | ira | *they* |
| 194. | *si(dr)i(t) | ?si:rI | siri | *semen, masturbation* |
| 195. | *sikita | si:xE | sixa | *enmity, hostility* |
| 196. | *siku(n) | ʊ́:xʊ̄ | ʊ́xʊ́ | *tail, tail of fish* |
| 197. | *nsila(k) | ttʊ́lE | ttʊ́la | *lightning, to shine* |
| 198. | *sili | tilI | tili | *to enter* |
| 199. | *sina(R) | ttʊ́lE | ttʊ́la | *to shine* |
| 200. | *nsipo | tiwE | tiwe | *downwards* |
| 201. | *nsiwa | ti:wE | tiwa | *nine* |
| 202. | *so(dr)i | sorou (old form) torou (new form) | sorou, torou | *humility toward a chief* |
| 203. | *nsoka | tokatok0 | toka-toka | *to stab* |
| 204. | *sola | to:l0 | tola | *coconut blossom* |
| 205. | *suku | ?tʊ́:tʊ́ | tʊ́ʊ́-tʊ́ʊ́ | *to bathe, dive, wash* |
| 206. | *sulu | ttʊ́lE | ttʊ́la | *torch, glow* |
| 207. | *susu | tʊ́:tʊ̄ | tʊ́tʊ́ | *breast, suck* |
| 208. | *susu(dr) | ti:tI | titi | *to sew* |
| 209. | *-nta | -šE | -ša | *our* (incl.) |
| 210. | *ta?aki | sa:xi | saaxii *(extract it)* | *to draw water, dig up* |
| 211. | *ta-dawa | xa:rawerawE | xaa-rawarawa | *green* |
| 212. | *tali | ta:lI | tali | *cord, rope* |

| | POC | WOL surface form | WOL base form | Gloss |
|---|---|---|---|---|
| 213. | *taliŋa | taliŋE | taliŋa | *ear, earwax* |
| 214. | *ntalo | ?ccelI | cceli | *shore tree; Calophyllum inophyllum* |
| 215. | *tama | ta:mE | tama | *father* |
| 216. | *tamole | mwa:lE<br>tamwelʊ̈ *(chief)* | mwale,<br>tamwelʊ̈ | *man* |
| 217. | *tano(?) | ta:lO | talo | *earth, soil* |
| 218. | *taŋi(s) | taŋI | taŋi | *to cry, weep* |
| 219. | D.*tampi | ?tapiyE | tapiya | *bowl* |
| 220. | *tampu | ta:bU | tabu | *a ban, taboo* |
| 221. | *tapu- | tawi | tawii | *conch* |
| 222. | *tansi(k) | ta:tI | tati | *sea, salt water* |
| 223. | D.*tasimi | taimi | taimii | *sharpen it* |
| 224. | *tau- | ?tau- *(practitioner)* | tau- | *man, person* |
| 225. | *tia(n) | siyalE | siala | *belly* |
| 226. | *tido | surO | suro | *to look at* |
| 227. | *tika | sixE *(angry)* | sixa | *bad* |
| 228. | *tina | si:lE | sila | *mother* |
| 229. | *tipi | si:fI | sifi | *man's girdle, woman's skirt* |
| 230. | *tiRi | si:rI *(masturbate)* | siri | *to spurt, urine* |
| 231. | *tiRo(m) | ?sa:rʊ̈ | sarʊ́ | *oyster* |
| 232. | *toka | toxO | toxo | *to arrive, land* |
| 233. | *toko(n) | so:xO | soxo | *staff, pole* |
| 234. | *tolu | seli- | seli- | *three* |
| 235. | *tom(i) | ?sorom(i) | sorom(ii) | *to drink, sip* |
| 236. | *topu | ?wo:u | ou | *sugar cane* |
| 237. | *tu?a | so:wE | soa | *back, beyond, outside, edge* |
| 238. | *tu(dr)(i) | šʊ́ | šʊ́ʊ́ | *bone, body* |
| 239. | *tuki | sʊ́xʊ́ *(hit it)* | sʊ́xʊ́ʊ́ | *to hammer, pound* |
| 240. | *tupa | su:pO | supa | *fish poison* |
| 241. | *tumpu(?) | su:bU | subu | *to be born* |
| 242. | *turu | sʊ́:rʊ̈ | sʊ́rʊ́ | *post, kneel, knee* |
| 243. | *tusu(k) | ?ti, xati:ti | tii,<br>xa-tii-tii | *to point, index finger* |
| 244. | *tu?u(d) | sʊ́ | sʊ́ʊ́ | *to stand up* |
| 245. | *?uda(ŋ) | ʊ́:rE | ʊ́ra | *lobster* |
| 246. | *?uluŋa | ilʊ́ŋE | ilʊ́ŋa | *pillow* |
| 247. | *?umu | u:mwU | umwu | *earth oven, to roast* |
| 248. | *?una(p) (D.*?una) | ʊ́:lE | ʊ́la | *fish-scale, body hair* |
| 249. | *upe | ʊ́:fE | ʊ́fa | *seedling, seeds* |
| 250. | *uRita | xʊ́:sE | xʊ́sa | *octopus* (cf. *kuRita) |
| 251. | *?utup | itiitI | iti-iti | *to flood, draw water* |
| 252. | *waka | wexarE | waxara | *root* |
| 253. | *waŋka(ŋ) | wa | waa | *canoe* |
| 254. | *walu | wa:lI | wali | *eight* |
| 255. | *waRo | yɔ | yɔɔ | *string, rope* |
| 256. | *wasa | ?ta:wE | tawa | *open sea* |
| 257. | *yaŋo | yaŋoyaŋO | yaŋo-yaŋo | *yellow* (cf. *aŋoaŋo) |

## NOTES

1. In the writing of this paper, the labour was divided in such a way that Tawerilmang provided the Woleaian (WOL) linguistic data, while Sohn was responsible for the analysis. This paper benefited from a comparative Micronesian seminar conducted by Dr. George Grace in 1972-73.

For a synchronic description of the phonology and syntax of WOL, see Sohn 1975.

2. Other POC reconstructions from Grace n.d. a and n.d. b are occasionally cited. These are prefixed with a D. (referring to Grace's source, Dyen 1949).

3. Lee (1976) makes a strong case for certain kinds of rule ordering in natural phonology.

4. Judging from the description of Marshallese stress in Bender (1975), dissimilatory a-raising may be related to alternating stresses in Woleaian. We must defer the study of WOL stress for the future, however.

5. Base forms of some more thematic-stem transitive verbs are given below:

| | | | | |
|---|---|---|---|---|
| baisii | *untie it* | (cf. | bai-bai | *to untie*) |
| bboolii | *pound it* | (cf. | bboo | *to pound*) |
| beli ii | *snap it* | (cf. | beli-beli | *to snap off*) |
| fatoxii | *plant it* | (cf. | fato-fato | *to plant*) |
| filetii | *stir it* | (cf. | file-file | *to stir*) |
| narii | *taste it* | (cf. | na-na | *to taste*) |
| raxomii | *hug it* | (cf. | rraxo | *to hug*) |
| tɔɔfii | *rub it* | (cf. | tɔɔ-tɔɔ | *to rub*) |

6. This appears to be a case of idiosyncratic final-vowel shortening, i.e. ʉʉ > ʉ. This exception may be explained away by reconstructing *pitu? for some pre-WOL (or Proto-TK) stage.

## BIBLIOGRAPHY

BENDER, Byron W.

1973 Parallelisms in the morphophonemics of several Micronesian languages. *WPLUH* 5/8:1-16. Reprinted in *Oceanic Linguistics* 12 (1973):455-477.

1975 Stress in Marshallese. MS. University of Hawaii.

CHOMSKY, Noam and Morris HALLE

1968 *The sound pattern of English.* New York: Harper & Row.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

GRACE, George

1969 A Proto-Oceanic finder list. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

n.d.a. On the history of the Trukese sound system. MS. University of Hawaii.

n.d.b. Some Trukese cognates (from Dyen 1949). Mimeo. University of Hawaii.

HOOPER, Joan B.

1974 Aspects of natural generative phonology. Doctoral dissertation, University of California at Los Angeles. Reproduced by Indiana Linguistic Club, Bloomington, Indiana.

LEE, Gregory

1976 Rule limitation in phonology. MS.

SOHN, Ho-min

1971 a-raising in Woleaian. *Working Papers in Linguistic, University of Hawaii* 3/8:15-35.

1975 *Woleaian Reference grammar.* Honolulu: University Press of Hawaii.

SOHN, Ho-min, Anthony TAWERILMANG, Isaac LANGAL, and Celestine YANGILMAU

1977 Consonant shifts and subgrouping in the Sonsorol-Ulithi-Woleai chain. *Language Sciences* 44:19-24.

# SURFACE AND DEEP CONTRASTS IN ULITHIAN PHONOLOGY

Ho-min Sohn and Byron W. Bender

One of the major problems in Ulithian (ULI) phonology is that there are many asymmetrical and limited phonetic contrasts. The aim of this paper is to present some conspicuous contrast patterns of such limited distribution and associated problems in Section 1, to suggest solutions to them in Section 2, to discuss a problem concerning the third person singular object suffix forms in Section 3, and to illustrate underlying contrasts in Section 4.

As will be seen, we propose to set up *base forms* of morphemes and words, which have been decided upon in light of the general phonological characteristics observable in various morphophonemic changes in the surface forms of ULI.[1] Such phonological characteristics are mostly synchronic, but occasionally diachronic evidence is taken into account in cases of indeterminacy.

There are several definite advantages to this approach in which phonemicisation is determined by base form contrasts. For example, various asymmetrical and limited phonetic contrasts can be accounted for in terms of environmental conditioning; the alternation of stem-final vowels may be explained by general phonetic rules; and the maximum uniformity thereby attained in base forms facilitates syntactic description, because a given grammatical or lexical item generally has one and only one phonemic shape.

## 1. SURFACE FORM CONTRASTS

1.1. A superficial non-contrast might be noticed between the velarised bilabial stop mw and its plain counterpart m when they are unreleased, i.e. before a pause or a homorganic consonant. Consider, however, the following pair:

(1) a. [lɑ:ʲm] *mosquito*
b. [la:m] *lamp; clear*

Spectrographic analysis indicates that the second format of the vowel in (1a) falls rapidly to a back [u] or [o] position, showing the velarisation of the following consonant, while the vowel in (1b) shows a smooth transition to a plain [m] position. This evidence and the distributional limitation of [ɑ(:)] have led us to the conclusion that the significant contrast is between the two final nasals, and that the backing of the vowel [ɑ:] in (1a) is caused by the [+back] feature of the final nasal.

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 201-213.
*Pacific Linguistics*, C-80, 1984.

1.2. Two [ l ]'s of different quality contrast in certain limited environments. In the first place, the so-called 'construct suffix' morpheme [ l ] *of* (referred to as Cs-ℓ) and the third person singular possessive suffix morpheme [ l ] *his* (3s-ℓ) contrast in many instances. In such contrasts, Cs-ℓ is light and relatively fronted, while 3s-ℓ is dark and retracted. The different qualities of the two [ l ]'s may be reflected in preceding stem vowels (see below), which fact might lead one to suppose that the grammatical difference between the two morphemes concerned is manifested by the stem vowels, and that there are different allophones of l following these vowels.[2] Consider for example the following:

| | | | | | |
|---|---|---|---|---|---|
| (2) | [ubʷel<] | *chest of* | : | [ubʷɑl>] | *his chest* |
| | [tapal<] | *cheek of* | : | [tɑpɑl>] | *his cheek* |
| | [sogoˆl<] | *stick of* | : | [sogoˇl>] | *his stick* |

No contrasts are found if the vowel preceding [ l ] is a single (short) [ i ], [ e ], [ u ], or [ ə ]. However, long vowel [ e: ] does show a contrast before the two suffixes.

| | | | | | |
|---|---|---|---|---|---|
| (3) | [ŋi:l<] | *teeth of* | : | [ŋi:l<] | *his teeth* |
| | [raxel<] | *age of* | : | [raxel<] | *his age* |
| | [lurul<] | *shade of* | : | [lurul<] | *his shade* |
| | [ɓʷɔ:dəl<] | *nose of* | : | [ɓʷɔ:dəl<] | *his nose* |
| | [pece:ˆl<] | *foot of* | : | [pece:ˇl<] | *his foot* |

Regarding the [ l ]'s of the two morphemes as phonemically identical would require one to set up a number of new vowel phonemes, e.g. a in contrast with ɑ, e in contrast with eˇ, etc. Besides, no regularity could be captured in the alternation between the two series of stem vowels:

| | | | | | | |
|---|---|---|---|---|---|---|
| (4) | [e] | + | Cs-ℓ | vs. | [ɑ] | + 3s-ℓ |
| | [e] | + | Cs-ℓ | vs. | [e] | + 3s-ℓ |
| | [a] | + | Cs-ℓ | vs. | [ɑ] | + 3s-ℓ, etc. |

As an alternative to such a treatment, which would set up a number of additional vowel phonemes of limited distribution, two distinct ℓ phonemes might be postulated, with the vowel contrasts accounted for as allophonic variants conditioned by the l's. But the two l's would actually contrast only in these two morphemes (Cs and 3s) after certain vowels; in all other positions (with the possible exception of (6) below) they would be in complementary distribution, varying mechanically according to the neighbouring vowels. One could emphasise this limitation in a transcription by using an l archiphoneme in the non-contrastive positions — using, for example, l's with fronting and backing diacritics for the fully specified l's, and an l without any diacritic for the archiphoneme.

| | | | | |
|---|---|---|---|---|
| (5) | wubʷel< | *chest of* | wubʷal> | *his chest* |
| | tapal< | *cheek of* | tapal> | *his cheek* |
| | raxel | *age of* | raxel | *his age* |

This solution implies the existence of two allomorphs for each of the two suffix morphemes. Or one could follow the principle 'once a phoneme, always a phoneme' (Householder 1964:25) and extend the contrast to the many non-contrastive positions in which ℓ may occur, but this would introduce a great deal of unnecessary redundancy and tend to obscure the severely limited distribution of the l-contrast. None of the solutions discussed thus far would provide a basis for any simple generalisation concerning the vowel alternations before the two suffixes in some forms but not in others.

A second instance of the l-contrast may be found in the following examples:

(6) (a) [xaɗal<+bʷo] *to walk in line because*
(b) [xaɗal>bʷo] *supposedly*

(c) [xaƀʷl<e] *we* (excl.) *will...*
(d) [xaƀʷl>e] *to miss him*

(6a) and (6b) differ from each other in several respects: (i) the quality of the [l]'s, (ii) the quality of the low vowels, and (iii) the presence of a pause juncture (+) in (6a). (6c) and (6d), on the other hand, differ only in the quality of the [l]'s. This suggests that the difference in the low vowels preceding the different [l]'s in (6a) and (6b) is conditioned by the latter, rather than the other way around. There being no indication that the quality of l is determined by the presence or absence of pause, one would have to conclude from the evidence of the forms in (6) that there would seem to be no other solution than to set up two underlying l's.

**1.3.** In word-final position, there is a contrast between a sequence consisting of the voiceless [x] followed by a [+back] voiceless vowel (i.e. [x̥, ɑ̥, ʊ̥]), the same [x] followed by Ø, and the sequence of the voiced velar fricative [ɣ] followed by a corresponding voiced vowel:

| | | | | | |
|---|---|---|---|---|---|
| (7) | [ƀʷuraxɑ̥] | *smoke* | : | [maleka:xɑ̥] | *travel* |
| | [xa:xɑ̥] | *helm* | : | [ɗoxɑ̥] | *to get* |
| | [kəlɔxɑ̥] | *hungry* | : | [ƀʷarəxɑ̥] | *dance* |
| | [ƀʷulaxɑ̥] | *taro* | : | [fasamaxɑ̥] | *pebble* |
| | [maŋa:xʊ̥] | *clothes* | : | [ƀʷuxʊ̥] | *knot* |
| | [tet:eraχ] | *sailing* | : | [ƀʷareχ] | *pain* |
| | [ləƀʷa:χ] | *to hide* | : | [tɔgɔχ] | *coconut shell* |
| | [kaɣo] | *box* | : | [tamaɣo] | *cigarette* |

The most conspicuous occurrence of this kind is associated with the second person singular object suffix, which has the phonetic form [xɑ̥] and its allomorph [xʊ̥]. The latter occurs after a high vowel (e.g. i or u) and the former elsewhere.

(8) [pa:luxʊ̥] *to lead you*
cf. [pa:liyei] *to lead me*
[tavaxɑ̥] *to cut you*
cf. [tavayei] *to cut me*
[ɗaƀʷoxɑ̥] *to follow you*
cf. [daƀʷeyei] *to follow me*

The above surface information might lead one to postulate a set of voiceless vowels of extremely limited distribution (finally after x), or else a ɣ phoneme (in partial complementation with x) to account for the final vowels that do not devoice.

**1.4.** There is a contrast of high frequency of occurrence between the two low vowels [a] and [ɑ]. They contrast not only before the two morphemic [l]'s (**Cs** -ℓ and **3s** -ℓ) as in (9) but also in many independent forms as in (10).

| | | | | | | |
|---|---|---|---|---|---|---|
| (9) | short | [wagal<] | *vein of* | : | [wagal>] | *his vein* |
| | | [tal<] | *rope of* | : | [tal>] | *its rope* |
| | | [tapal<] | *cheek of* | : | [tapal>] | *his cheek* |
| | long | [wa:l<] | *canoe of* | : | [wa:l>] | *his canoe* |
| | | [ifa:l<] | *underside of*: | | [ifa:l>] | *its underside* |
| (10) | short | [maθ] | *sated* | : | [maθ] | *cooked* |
| | | [ɓʷal<] | *to inspect* | : | [ɓʷal>] | *stuck* |
| | long | [ya:f] | *fire* | : | [ya:f] | *swimming* |
| | | [ta:l<] | *rope* | : | [ta:l>] | *well versed* |
| | | [fa:s] | *stone* | : | [fa:s] | *penis* |

In spite of the high frequency of the contrast, several points may be noted against the establishment of the two vowels as separate phonemes. In the first place, there is no contrast in the environment C_#. Observe the following examples of low and front vowel contrasts:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (11) | [ri] | *spouse* | [ɬi] | *to kill* | [ɓʷi] | *sides of vagina* |
| | [re] | *they* | [ɬe] | *this* | [ɓʷe] | *fortune-telling* |
| | [ræ] | *side* | [ɬæ] | *as* | [ɓʷæ] | *to find out* |
| | [ra] | *branch* | [ɬa] | *that* | [ɓʷa] | *rotten* |
| | * [ra] | | * [ɬa] | | * [ɓʷa] | |

Secondly, [a] in C_# is differentiated into front and retracted varieties when the Cs-ɬ and 3s-ɬ suffixes follow, which suggests that the contrast may not be inherent but conditioned by the suffixal environment.

(12)

[ra] *branch* → [ra:l>] *branch of*

[ra] *branch* → [ra:l>] *its branch*

Thirdly, no contrast has been found in the environment X__Y in which X is non-null and Y is a syllable, though both the fronted and retracted low vowels may occur in this position. That is, their occurrences are perfectly predictable according to X and Y. Observe that the quality of the first vowel in each word in (13) can be predicted from the vowel in the following syllable, i.e. fronted if the following vowel is one classed as non-back (i.e. [i, e, æ, ə, a]), and retracted elsewhere.

| | | |
|---|---|---|
| (13) | [mədep] | *a part* |
| | [madar] | *to disperse* |
| | [mədel<] | *being sated of* |
| | [madal>] | *its being cooked* |

In short, the contrast between [a] and [a] is limited to the environment __C#.

1.5. A final problem regarding surface contrasts is raised by the existence of two different monophthongal vowel qualities in the mid-central area that contrast only when long. For the moment, the two qualities are transcribed as [ə] and [ə̈]:

| | | | | | |
|---|---|---|---|---|---|
| (14) | [tə:s] | *truth* | : | [të:s] | *porcupine fish* |
| | [ɖə:r] | *lavalava* | : | [ɖë:r] | *indebtedness* |
| | [pə:l] | *emptiness of* | : | [pë:l] | *arm of* |

The vowel in the first column is common, while the vowel in the second is rare but also occurs in a few other forms such as:

| | | |
|---|---|---|
| (15) | [yë:r] | *accustomed; name of an island* |
| | [më:l] | *name of a star* |
| | [më:r] | *fresh (tree, vegetables)* |
| | [cë:l] | *leaf of* |

Sound spectrographs show that both vowels are monophthongal, and that F1 of the vowel in [tə:s] = 450 and F2 = 1600; F1 of the vowel in [të:s] = 550 and F2 = 1400. The above formants indicate that the two sounds [ə] and [ë] are very close to the French vowels Ø and œ respectively (Delattre et al. 1952:198). [ə] seems to have more lip rounding than [ë].

Several problems are involved in the phonemic interpretation of the two sounds. First of all, the number of forms containing [ë] is small. Secondly, the contrast between the two is limited to long segments. Thirdly, [ë:] in C__C is not shortened before a suffix, while [ə:] may be:

| | | | | | |
|---|---|---|---|---|---|
| (16) | [ɖə:r] | *lavalava* | : | [ɖərəl] | *lavalava of* |
| | [ɖë:r] | *indebtedness* | : | [ɖë:rəl] | *indebtedness of* |
| | [tə:s] | *truth* | : | [tə:səl] | *truth of* |
| | [të:s] | *porcupine fish* | : | [të:səl] | *porcupine fish of* |

These limitations make one hesitant to set up two mid-central vowel phonemes.

## 2. SUGGESTED SOLUTIONS

2.1. Most of the problems raised in the preceding section can be solved simply, straightforwardly, and with greater generality by introducing non-*ad hoc* base forms in the phonology. Then, as will be seen, superficial contrasts of limited distribution turn out to be allophonic variants conditioned by material present in the base forms but not readily observable on the surface, or the result of geminates versus non-identical vowel qualities. Such base and surface forms will be directly related by a series of ordered phonological rules. The approach followed here not only solves the problem of limited distribution but also contributes to regularising most of the irregular morphophonemic alternations.

2.2. A solution to the problem of surface vowel contrasts before bilabial nasals (see 1.1.) has already been implied in the earlier discussion; the consonants can be said to contrast as mw and m in base forms, conditioning vowel allophones that in some cases constitute their chief stigmata. Thus, for example, the base forms of the items in (1) are established in (17).

| | | | | | |
|---|---|---|---|---|---|
| (17) | a. | lamwo | *mosquito* | (cf. [lɑmʷol>] | *mosquito of* ) |
| | b. | lama | *lamp; clear* | (cf. [lamal<] | *lamp of*) |

Phonological rules will impose the [+back] feature of mw onto the preceding vowel, lengthen the non-final vowels, and delete the final vowels.

With regard to the problem of two [l]'s of different quality, the Cs and 3s morphemes may be set up as li and la respectively. Seemingly contrasting stem vowels before the two morphemes can then both be derived from the same basic stem vowel as the result of the conditioning of the two contrasting basic vowels i and a in the suffix morphemes -li and -la. Although -li and -la lost their final vowels, as do all other lexical base forms, they still carry features of the vowels such as [+high, -back, -low] from i, and [-high, +back, +low] from a, which cause the alternations of the stem vowels. Thus the process is active, i.e. synchronic. The fronted and retracted l's thereby conditioned do not affect the basic stem vowels i, u, ə, or short e. Rather, the *l*'s are assimilated in quality to these mid or high vowels. These four basic stem vowels happen to correspond to Sonsorolese (SNS) high vowels in most instances (Bender 1967 and Quackenbush 1968).[3]

| (18) | ULI base forms | | SNS surface forms | |
|---|---|---|---|---|
| | taŋi | [taŋ] | [taŋi̥] | *to cry* |
| | cf. [taŋil<] (+ Cs) | | | |
| | lutu | [lut] | [rutu̥] | *to jump* |
| | cf. [lutul<] (+ Cs) | | | |
| | maulu | [maul] | [maru̥] | *war* |
| | cf. [maulul<] (+ Cs) | | | |
| | yaɗe | [ya:θ] | [ya:ti̥] | *gall bladder* |
| | cf. [yaɗel<] (+ Cs) | | | |
| | taxuru | [tɑgur] | [tal̈ɨxi̥] | *back* |
| | cf. [tagurəl<] (+ Cs) | | | |
| | faɗə | [fa:θ] | [fa:ti̥] | *eyebrow* |
| | cf. [faɗəl<] (+ Cs) | | | |

From this comparative evidence, we tentatively conclude that only those vowels which were historically high vowels are resistant to the influence from the following l's of different quality. This may be supported by the fact that SNS (and also Woleaian (WOL)) preserves many reflexes of original forms. For WOL reflexes, see Tawerilmang and Sohn 1976.

The irregular alternation of the stem vowels before the two morphemes Cs and 3s, as illustrated in (4), turns out to be regular if we set up well-motivated basic stem vowels covering all relevant lexical items and develop an appropriate set of morphophonemic rules to map the base forms onto surface manifestations. Thus, for example, the following are proposed as the basic stem vowels underlying (4):

(19)

| | | | |
|---|---|---|---|
| a | [e] | : | [ɑ] |
| | [a] | : | [ɑ] |
| e | [e] | : | [e] |

The difference in alternation between [e] : [ɑ] and [a] : [ɑ] is due to the influence of the vowel in the preceding syllable, e.g. [e] : [ɑ] only when the preceding vowel is high.[4] (20) gives the base forms corresponding to (2) and (3).

| (20) | | | | | |
|---|---|---|---|---|---|
| | uba-li | *chest of* | : | uba-la | *his chest* |
| | tapa-li | *cheek of* | : | tapa-la | *his cheek* |
| | soxo-li | *stick of* | : | soxo-la | *his stick* |
| | ŋii-li | *teeth of* | : | ŋii-la | *his teeth* |
| | raxe-li | *age of* | : | raxe-la | *his age* |
| | luru-li | *shade of* | : | luru-la | *his shade* |
| | bɔɔɖə-li | *nose of* | : | bɔɔɖə-la | *his nose* |
| | pecee-li | *foot of* | : | pecee-la | *his foot* |

The reconstruction of base forms for stems as well as for -li and -la is synchronic but largely corresponds to the forms of PMC (Proto-Micronesian), WOL, and SNS. Thus the reconstructed stem vowels not only have historical and comparative significance, but they also indicate the traditional declensional classes to which the nouns belong. A sample set of ULI, WOL, and SNS forms follows, in which the similarity between the ULI base forms and the corresponding SNS independent forms should be noted.[5]

| (21) | base form | ind. form | + Cs | + 3s |
|---|---|---|---|---|
| ULI | yafara | [yɑfɑr] | [yafaral<] | [yɑfɑral>] |
| WOL | -- | [yefarę̥] | [yeralį̥] | [yeralę̥] |
| SNS | -- | [yafalą̥] | [yafalarį̥] | [yafalarą̥] |
| | *shoulder* | | | |
| ULI | imwa | [i:$^{i}$m] | [imʷel<] | [imʷɑl>] |
| WOL | -- | [i:mʷę̥] | [imʷelį̥] | [imʷalę̥] |
| SNS | -- | [i:mʷą̥] | [imʷerį̥] | [imʷarą̥] |
| | *house* | | | |

The second instance of contrasting ℓ's (6) can also be resolved by setting up proper base forms. Again, the two ℓ's are allophones of a single underlying l. The processes involved are as follows:

(22) base

| | | | |
|---|---|---|---|
| | xaɖale#bo | => xaɖa<l<e#bo | |
| => | xaɖa<l<#bo | => xa>ɖa<l<+ bo | |
| | | [xaɖal<+bʷo] | *to walk in line because.....* |
| | xaɖalboo | => xaɖal>bo | |
| | | [xɑɖɑl>bʷo] | *supposedly* |
| | xa#be#le | => xa#be#l<e | |
| => | xabel<e | => xaɓl<e | |
| | | [xɑɓʷl<e] | *we* (excl.) *will* |
| | xabɔle + ya | => xabɔley | |
| => | xabɔle | => xable | |
| | | [xɑɓʷl>e] | *to miss him* |

2.3. The problematic contrasts among final-syllable velar fricatives and their following vowel (or lack thereof) (see (7) and (8)) can be resolved most efficiently by setting up base forms in such a way that surface voiceless vowels are represented as [+back] single vowels and surface voiced vowels as geminates, while Ø vowels which in this position are [-back] are set up in the same way as are all other final vowels that delete. This treatment is well motivated in that (a) no words end in a double vowel on the surface, and (b) when suffixes

are added, voiceless vowels and Ø vowels are realised as single voiced vowels, while voiced vowels are realised as geminate vowels. These two points are basic for the establishing of all ULI base forms. Thus the phonetic manifestations of independent base forms are effected by simply devoicing single [+back] vowels after x but dropping [-back] vowels in word-final position, including the position x__#. The examples in (7) can be phonemicised as the following base forms. The corresponding phonetic forms followed by Cs-ℓ are also given for reference.

| | | | |
|---|---|---|---|
| (23) | buraxo | *smoke* | [bʷuraɑgol<] |
| | malekaaxo | *travel* | [malekɑ:gol<] |
| | xaaxo | *helm* | [xɑ:gol<] |
| | ɗoxo | *to get* | [ɗogol<] |
| | kəlɔxo | *hungry* | [kəlɔgol<] |
| | barəxo | *dance* | [bʷarəgol<] |
| | bulaxa | *taro* | [bʷulagal<] |
| | fasamaxa | *pebble* | [fasamagal<] |
| | maŋaaxu | *clothes* | [maŋɑ:gul<] |
| | buxu | *knot* | [bʷugul<] |
| | tetteraxe | *sailing* | [tet:eragel<] |
| | bbarexe | *pain* | [bʷ:aregel<] |
| | ləbaaxe | *to hide* | [ləbʷa:gel<] |
| | tɔxɔxə | *coconut shell* | [tɔgɔgəl<] |
| | kaxoo | *box* | [kɑgo:l<] |
| | tamaxoo | *cigarette* | [tɑmɑgo:l<] |

In the same way, the examples in (8) are the surface forms derived from the base forms in (24). Notice the morphophonemic changes effected in the environment of the morpheme boundaries.

| | | | | | |
|---|---|---|---|---|---|
| (24) | paali + yei | *to lead me* | : | paali + xo | *to lead you* |
| | tafa + yei | *to cut me* | : | tafa + xo | *to cut you* |
| | ɗabe + yei | *to follow me* | : | ɗabe + xo | *to follow you* |

2.4. That the low vowel qualities [a] and [ɑ] turn out to be predictable within the general framework adopted here should come as no surprise in light of the previous discussions, which indicated that the contrast between them was limited to the environment __C#, and that certain contrasts in this environment could be conditioned by deleted final stem vowels set up following the consonant in base forms. Thus, [a] and [ɑ] prove to be conditioned variants of a. Some of the examples in (9) and (10) are phonemicised as (25).

| | | | | |
|---|---|---|---|---|
| (25) | waxa-li | *vein of* | waxa-la | *his vein* |
| | waa-li | *canoe of* | waa-la | *his canoe* |
| | maɗe | *sated* | maɗa | *cooked* |
| | yafe | *fire* | yafa | *swimming* |

2.5. With regard to the two mid-central vowel contrasts, the problem lies in the interpretation of [ë:]. It turns out that words containing this sound correspond to words that have a vowel cluster rather than a long monophthongal vowel in other TK languages.

(26) To ULI [pë:] *arm* and [pë:l] *arm of* correspond:
SNS: paai and paɨrI
Tobi (TBI): paa and paɨr
WOL: paaʊ̈ and paʊ̈lI
Ifaluk: paa and pawul

To ULI [ca:y] *leaf* and [cë:l] *leaf of* correspond:
SNS: saaɨ and saaɨrI
TBI: caaɨ and caaɨr
WOL: šəə and šəəlI

To ULI [të:s]
SNS: taɨθɨ
TBI: taɨs

We are reluctant to establish [ë:] as a new phoneme for the distributional reasons indicated in the preceding section, and this comparative evidence suggests that we may be able to interpret it as a cluster of already established phonemes instead. Internal evidence such as the alternation between the independent and construct forms of 'leaf' also point in this same direction. For this item the base form can be established as cayə and then the independent form will be derived by dropping the final vowel and lengthening the remaining vowel, i.e., [ca:y].

[pë:] and [pë:l] should be handled differently because they do not show parallel alternation. ULI æ corresponds in many cases to a of other TK languages, and ULI u to ɨ. Thus the base form of [pë:] will be set up as pæu. (We cannot set it up as pau because of the existence of forms such as [mɑul] *war*, for which the basic cluster au is needed). The other examples such as [të:s] *porcupine fish* and [ɗë:r] *indebtedness* may be handled in the same way as [pë:]. A later phonetic rule should specify the phonetic quality of these phoneme composites.

## 3. THIRD PERSON OBJECT SUFFIXES

One matter particularly difficult to decide upon was the base form of the third person singular object suffix, which corresponds to the underlined surface forms in the following examples.

(27) [lli-(y)] *kill him*
[fə:ru-(y)] *make it*
[fɑgo-(y)] *miss him*
[xɑmɑɗɑvɑ-Ø] *explain it*
[ɗabwe-(y)] *follow him*

That is, the suffix is realised on the surface as an optional [y] except for the position after [a] where it is zero. There is some evidence, however, in support of the assumption that the base form of the suffix is -ya.[6]

Although ya never appears phonetically as [yV] when the verb to which it is suffixed functions as the main verb, it does appear as such if the verb is nominalised before an attributive suffix.

(28) [lliyei] *what I killed*
[lliyəmʷ] *what you killed*
[lliyal>] *what he killed*
[lliyel<se mal<] *what someone killed*
[lliyac] *what we* (incl.) *killed*

| | |
|---|---|
| [lliyæ:r] | *what they killed* |
| [xamaɗavayey] | *what I explained* |
| [xamaɗavayəmʷ] | *what you explained* |
| [xɑmaɗɑvɑyɑl>] | *what he explained* |
| [xamaɗavayel<se mal<] | *what someone explained* |
| [xɑmaɗɑvɑyɑc] | *what we (incl.) explained* |
| [xamaɗavayæ:r] | *what they explained* |

If in (28) the forms (e.g. [ye], [yə] and [yɑ]) between a verb stem (i.e. [lli], [xɑmaɗɑvɑ]) and an attributive suffix (e.g. [y], [mʷ], [l>], [l<], [c]) are reduced to the base form -ya, then general morphophonemic rules will derive all the forms related to the base -ya unambiguously and without exception.

Exactly the same morphophonemic behaviour may be observed in many words which are not transitive verbs. For example, [fɑlu(y)] *island* and [cu(y)] *to disappear* retain the optional [y] in their independent form, and if attributive suffixes are added, the following phonetic forms are obtained:

(29)

| | |
|---|---|
| [fɑluyey] | *my island* |
| [fɑluyəmʷ] | *your island* |
| [fɑluyɑl>] | *his island* |
| [fɑluyel< se mal<] | *someone's island* |
| [fɑluyæ:r] | *their island* |
| [cuyey] | *my disappearance* |
| [cuyəmʷ] | *your disappearance* |
| [cuyɑl>] | *his disappearance* |
| [cuyel< se mal<] | *someone's disappearance* |
| [cuyæ:r] | *their disappearance* |

Since the base forms of [fɑlu(y)] and [cu(y)] have been set up as faluya and cuya, the parallelism between (28) and (29) leads to the reconstruction of the forms in (27) except for [xɑmaɗɑvɑ] as follows.

(30)

| | |
|---|---|
| lli-ya | *kill him* |
| fəəru-ya | *make it* |
| faxo-ya | *miss him* |
| ɗabe-ya | *follow him* |

Thus the final vowel deletion rule and the rule of optional deletion of semivowels may connect (30) to (27). On the other hand, [y] simply does not occur on the surface in the environment a __#.

## 4. EXAMPLES OF CONTRASTS IN BASE FORM SEGMENTS

Minimal pairs have not been found in sufficient number to make phoneme attestation easy, but the following examples may suffice to show the contrasts:

consonants

| | | | | |
|---|---|---|---|---|
| p | pare | *a kind of fruit* | tapa | *cheek* |
| t | tale | *rope* | tata | *a kind of fish* |
| c | cale | *water* | faca | *pandanus fruit* |
| k | kakka | *to carry* | piskaa | *spear* |
| b | baɗe | *scar* | taba | *taboo* |
| f | fase | *stone* | yafa | *swimming* |
| ɗ | ɗaa | *intestine* | faɗe | *string* |
| s | sare | *big knife* | fasa | *penis* |

| | | | | |
|---|---|---|---|---|
| x | xapi | *bottom, hip* | waxa | *vein* |
| mw | mwale | *man* | mweemwee | *to look for* |
| m | male | *animal, bird* | lama | *light bulb* |
| n | naanaa | *mummy* | sukuunu | *school* |
| ŋ | ŋaaŋe | *I* | faŋa | *to permit* |
| l | laŋe | *sky* | cale | *water* |
| r | rale | *day* | baro | *box* |
| y | yala | *sun* | xaya | *fish hook* |
| w | waa | *canoe* | wawa | *stick dance* |

**short** vs. long consonants

Defective are (i) quasi- native n which has the lowest frequency of occurrence; (ii) x whose doubling results in kk.

| | | | | |
|---|---|---|---|---|
| p : pp | pale | *dry* | ppale | *light* |
| | capi | *ancestor* | cappa | *turning over* |
| c : cc | caŋa | *short of reach* | ccaŋa | *skinny* |
| k : kk | makili | *sugar cane* | makkala | *comb* (Falalop) |
| b : bb | barexe | *hot (taste)* | bbarexe | *pain* |
| | bece | *hot* | bbece | *white* |
| | uba | *chest* | bubbu | *fish spec.* |
| f : ff | fisi | *star* | ffisi | *lightning* |
| | yafe | *fire* | yaffe | *land crab* |
| d : dd | dare | *to walk* | ddare | *to run* |
| s : ss | soŋo | *mangrove* | ssoŋo | *angry* |
| mw : mmw | mwɔlo | *desire* | mmwolo | *generous, kind* |
| | mwixilici | *miser* | mmwixi | *pepper* |
| m : mm | mata | *eye* | mmata | *to wake up* |
| | madare | *to disperse* | mmadare | *to burst* |
| ŋ : ŋŋ | ŋata | *hole* | ŋŋata | *hurry* |
| | fayelaŋa | *world* | faiŋŋa | *itchy feeling* |
| l : ll | loyo | *perfume* | lloyo | *wet* |
| r : rr | ro- | *all* | rro- | *string-bound bundle* |
| | raxe | *year, age* | rrai | *happy* |

vowels

| | | | | | | |
|---|---|---|---|---|---|---|
| i | lli | *to kill* | cibe | *scissors* | | |
| e | lee | *this* | pecee | *leg* | | |
| æ | læ | *as, which* | pæce | *sexual lust* | | |
| a | laa | *that* | capi | *lavalava* | paca | *tail* |
| ɔ | cɔɔ | *people* | llɔyo | *sweet* | bɔlo | *feather* |
| u | cuu | *to meet* | lloyo | *wet* | bolo | *soil* |
| ə | tə | *for a moment* | ləŋə | *ant* | bbələ | *filthy* |

short vs. long vowels

In the base form phonemicisation, many superficial length contrasts are suppressed. For example, [pix] *to play ball (v.)* and [pi:x] *ball (n.)* have no contrast in length in the base. Both are derived from the same base pixi, but the noun form is lengthened while the verb is not, and final vowels are dropped from both.

| | | | | |
|---|---|---|---|---|
| i : ii | cima | *head* | ciifeli | *nail* |
| e : ee | fedexe | *fight* | feefele | *woman* |
| æ : ææ | bæxi | *to float* (Vt) | xæætaa | *to do what* |
| a : aa | xaamami | *we* (excl.) | xaamaama | *to practice* |
| | fale | *food pounder* | faale | *cynical* |

| | | | | |
|---|---|---|---|---|
| ɔ : ɔɔ | bɔŋu | *feast* | bɔɔɗə | *nose* |
| o : oo | bolo | *soil* | booto | *boat* |
| u : uu | lutu | *to jump* | luutu | *soft wood* |
| | buru | *high tide* | ɗuuɗuu | *to bathe* |
| ə : əə | dərɑ | *woman's lavalava* | təəsə | *truth* |
| | fəlaŋa | *ashes* | iɾəətə | *village* |

semivowels (y and w)

| | |
|---|---|
| wayele | *aeroplane* |
| buyowe | *fish-trap* |
| wele | *strange* |

## NOTES

1. Some of the traditional principles of so-called taxonomic phonemics (see Chomsky 1964:75ff) such as biuniqueness and invariance are disregarded in the base form phonemicisation. On the other hand, concepts like contrast and complementary distribution are relevant only at the level of underlying base forms. Bender (1968) also attempted a rigorous base form phonemicisation of Marshallese, where skewed surface distributions are of a different sort.

2. Dyen (1965:33ff) seems to be following such a hypothesis in his description of a partially parallel phenomenon in Trukese.

3. The corresponding WOL stem vowels are also high, as pointed out and illustrated in Sohn and Bender (1973:Preface).

4. In ULI, WOL, and SNS, the basic vowel a is raised to e when it occurs between two high (basic) vowels (Bender 1973 and Sohn 1971 and 1975).

5. The SNS data were provided by Celestine Yangilmau and the WOL by Anthony Tawerilmang. For a description of WOL phonology, see Sohn (1975:Ch.1).

6. An alternative analysis would lead us to end up with the base form -a as the object suffix. In this case, the semivowel y must be introduced by a rule, as proposed for WOL in Tawerilmang and Sohn 1976.

## BIBLIOGRAPHY

BENDER, Byron W.

1967 Some Micronesian word lists. Mimeo. University of Hawaii.

1968 Marshallese phonology. *Oceanic Linguistics* 7/1:16-35.

1973 Parallelisms in the morphophonemics of several Micronesian languages. *Working Papers in Linguistics, University of Hawaii* 5/8:1-16. Reprinted in *Oceanic Linguistics* 12 (1973):455-477.

CHOMSKY, Noam

1964 *Current issues in linguistic theory.* Janua Linguarum 38. The Hague: Mouton.

DELATTRE, Pierre, Alvin M. LIBERMAN, Franklin S. COOPER, and Louis J. GERSTMAN

1952 An experimental study of the acoustic determinants of vowel colour: observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8:195-210.

DYEN, Isidore

1965 *A sketch of Trukese grammar.* Essay 4, New Haven: American Oriental Society.

HOUSEHOLDER, Fred W.

1964 Three dreams of Modern Greek phonology. In Robert Austerlitz, ed. *Papers in memory of George C. Pappageotes,* 17-27. Supplement to *Word* 20/3. Special Publication No. 5. New York: Linguistic Circle of New York.

1967 Distinctive features and phonetic features. In *To honor Roman Jakobson: essays on the occasion of his seventieth birthday,* vol.2:941-944. The Hague: Mouton.

QUACKENBUSH, Edward M.

1968 From Sonsorol to Truk: a dialect chain. Doctoral dissertation, University of Michigan, Ann Arbor. (University Microfilms, 69-2375.)

SOHN, Ho-min

1971 *a*-raising in Woleaian. *Working Papers in Linguistics, University of Hawaii* 3/8:15-35.

1975 *Woleian reference grammar.* Honolulu: University Press of Hawaii,

SOHN, Ho-min and Byron W. BENDER

1973 *A Ulithian grammar. Pacific Linguistics,* C-27.

TAWERILMANG, Anthony F. and Ho-min SOHN

1983 Proto-Oceanic reflexes in Woleaian. See pp.168-196 in this volume.

Sohn, H-M. and Bender, B.W. "Surface and deep contrasts in Ulithian phonology". In Bender, B.W. editor, *Studies in Micronesian linguistics*.
**C-80:201-213**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.201*

# AN ORTHOGRAPHIC DESIGN FOR WOLEAIAN

Ho-min Sohn

The following discussion is concerned mainly with orthographic recommendations for Woleaian (WOL), but I hope it will also contribute to the solution of similar problems for neighbouring languages including Ulithian (ULI), Satawalese (STW), and Lamotrek, and more importantly, to the design of a unified orthography for all of the Yap Outer Islands. The present discussions will cover (1) some basic concepts of an orthography, (2) the sound system of WOL, (3) the proposed alphabet, and (4) the proposed spelling conventions. The last two constitute the proposed orthography of WOL.[1]

At the end of this paper, an appendix, is provided to briefly present the orthography adopted in Sohn (1975) and Sohn and Tawerilmang (1976) since the recommendations in this paper were made.

## 1. BASIC CONCEPTS

(1) Speech and writing. The people of Woleai need a standard way of writing that everyone can use since writing is very important in our use of language, especially in our modern society where living becomes more complex day by day. The main role of a writing system is to represent speech, thus enabling communication across time and space. Thus, speaking and hearing correspond respectively to writing and reading.

Three typical types of writing systems are in use in the world today. These are alphabets, ideographies, and syllabaries. Most of our familiar writing systems, including that of English, are alphabets. If a writing system uses symbols or letters, each of which directly corresponds to an individual sound, and not to a word or word meaning, it is called an alphabet. Thus, for example, k in the alphabet of Eng. represents a sound and not a meaning. If a writing system uses symbols or characters each of which directly corresponds to a word or word meaning, it is called ideographic. Thus, for example, the Chinese character 我 [wǒ] represents the meaning *I*. If a writing system uses symbols or letters each of which corresponds neither to an individual sound nor to a word, but to a syllable, it is called a syllabary. Thus, for example, the Japanese letter カ represents the syllable [ka] which indicates neither one sound nor a meaning but a syllable consisting of a consonant and a vowel.

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 215-236.
*Pacific Linguistics*, C-80, 1984.

Ideographies have the grave disadvantage that there must be a great number of symbols, because every language has tens of thousands of words, and the users of such a writing system are subjected to a much heavier burden on their memories than are those whose languages are represented alphabetically. For example, a dictionary of Chinese includes more than 50,000 different symbols or characters. On the other hand, syllabaries are adequate only in such a language as Japanese where there are a relatively small number of different syllables. The number of syllables depends on the number of vowels and consonants and the type of syllable structure. For example, Jp. has only fifty-one basic letters representing all possible syllables in the language, because the number of vowels and consonants is very small and the syllable structure is of the form of a consonant only (C-type), a vowel only (V-type), or a consonant followed by a vowel (CV-type).

WOL has a syllable structure as simple as Jp., but has many more vowels and consonants. This fact makes it disadvantageous for Woleaians to have a syllabary. In fact, WOL was represented by various types of syllabaries in the past (Riesenberg and Kaneshiro 1960). These syllabaries apparently originated as ideographic pictures, but they were syllabaries in that the symbols thus created were used to represent any syllables of the same sound, regardless of the original meaning. It is important to note that in spite of their large number, the symbols given in Figure 26 in Riesenberg and Kaneshiro (278) are far from adequate to represent all WOL words and sentences correctly and unambiguously.

Alphabets have an advantage over the other systems, in that only a handful of letters are used to represent an infinite number of words and sentences in a language. Since the number is small (for example, Eng. has twenty-six basic letter symbols), it is easy to learn, write, and read the letters.

For this reason, we propose adopting an alphabetic writing system for the representation of the WOL language.

(2) Relation between sound and letter. An ideal alphabet writing system should serve not only for native speakers who already know the language but also for those who either partly know the language or want to learn or work on it. If an orthography is to be devised only for those who have native control of the language, we do not need any strict correspondence between sound and letter, because a loose correspondence can also give the ordinary speaker enough indication of what has been said so that he can supply the rest. In any language, a great deal can be omitted without a corresponding loss in the transmittal of messages from one person to another. For example, if we use the traditional spelling shuh to mean both *mountain* and *basket* in spite of the difference in pronunciation, there is not much difficulty in the conveyance of messages, since native speakers would pronounce them correctly if given the context. However, for non-native speakers who want to learn a language or are in the process of learning, or even for those native speakers who are not mature in the language, a strict correspondence between sound and letter is important, for it will make language learning easier and faster if they can learn how to pronounce each written word and sentence correctly. Strict sound-letter correspondence is also desirable not only for the benefit of native speakers who are in the initial stage of reading, but also because writing has to be a correct representation of speech. We could devise a totally separate set of phonetic symbols instead of alphabetic letters as in English, but if the alphabetic letters can be used as phonetic symbols, as they are, nothing could be more efficient.

Furthermore, strict sound-letter correspondence will contribute to the economy of language users' efforts. Loose correspondence forces speakers to memorise spellings of individual words with much effort, because their pronunciation does not automatically provide correct spellings, as is obvious to users of Eng. Finally, strict sound-letter correspondence will lead easily to the standardisation of writing, since it provides native speakers with a norm in writing their language, i.e. the norm 'Follow your pronunciation'.

I consider an orthography to be ideal, then, if it is useful for both mature speakers and language learners, and if it is subject to easy learning and correct reading and writing. One basic requirement for such an orthography is that sound and letter must match as unambiguously as possible. The WOL orthography proposed here is devised with such objectives in mind.

## 2. THE WOL SOUND SYSTEM

The most important prerequisite for an orthographic device based on an alphabet is an intensive linguistic analysis in order to reveal the sound system of the language. A poor knowledge of a sound system often leads one to devise an orthography which contains too many or too few symbols on the one hand, and includes useless spelling conventions or misses important sounds on the other. The spelling proposed by Smith (1951), which is an alphabetic writing, is an example of an orthography based on a poorly analysed sound system, in which Smith sets up too many letters and poor spelling conventions.

My intensive analysis of WOL, with the help of Tony Tawerilmang from Woleai, has revealed that WOL has fifteen short and eleven long significant consonants, six short and eight long significant vowels, and two significant semi-vowels, significant in the sense that any interchange of them in words will necessarily be accompanied by a change in meaning of the words involved. For example, the replacement of [ i ] with [ a ] in [ rixi̥ ] *run* and [ raxi̥ ] *year* results in the change of meaning. (The square brackets mean the actual pronunciation of the forms. For the pronunciation of each symbol see the pages following. The small circle under a symbol indicates voicelessness of the sound.)

There are many non-significant sounds also, non-significant in the sense that they have no power to change meanings even if they were interchanged. Their differences, entirely automatic, are caused by their sound environments. For example, in [ rixi̥ ], exchange of the two sounds [ i ] and [ i̥ ], even if it were made, would not bring about any change in the meaning of the word. Their difference is quite automatic, [ i̥ ] appearing at word-final position and [ i ] between consonants or at word-initial position. Such non-significant sounds as [ i ] and [ i̥ ] are grouped into a significant sound /i/ for orthographic purposes. The significant sounds are technically called 'phonemes' and are transcribed in slashes.

### (1) WOL Consonants

WOL consonant phonemes are either single or double. The fifteen single consonants of WOL are classified as follows according to manner and place of articulation:

| manner of articulation | | | place of articulation | | | | |
|---|---|---|---|---|---|---|---|
| voice | mode | tongue shape | lips | lip-teeth | teeth-gum | hard palate | soft palate |
| voiceless | stop | plain | /p/ | | /t/ | /c/ | /k/ |
| | fricative | plain | | /f/ | /s/ | | /x/ |
| | | pulled back | /p̠/ | | | | |
| | | curled back | | | | /š/ | |
| voiced | | | | | | /r/ | |
| | nasal | plain | /m/ | | /n/ | | /ŋ/ |
| | | pulled back | /m̠/ | | | | |
| | flap | plain | | | /l/ | | |

The phonemes /p/, /t/, /c/, and /k/ are voiceless unaspirated stop consonants, pronounced approximately as p, t, ch, and k in the Eng. words 'spy', 'strong', 'matches', and 'sky' respectively. Single /c/ and /k/ occur only in loanwords. The phonemes /f/, /s/, and /x/ are voiceless fricative consonants, pronounced like f, s, and ch in Eng. 'five' and 'sun' and Ger. ich *I*, respectively. /x/ is voiced between two voiced vowels.

The phoneme /p̠/ is also a voiceless fricative consonant, produced as if one were blowing a candle while putting the tongue backward. Between two voiced vowels, it tends to be voiced. /š/ is another voiceless fricative, produced in the same way as the initial sound of Eng. 'shepherd' but with the tongue curled backward. /r/ which is a voiced fricative pronounced with the same tongue position as /š/. It is similar to the initial sound of Eng. 'read' except that for WOL /r/ the tongue is curled backward. While /š/ has a noticeable hissing sound, produced by squeezing the air through a narrow slit between the tongue tip and the palate, /r/ lacks such hissing because there is a wide groove between the tongue tip and the palate which allows the air to flow out freely.

The phonemes /m/, /n/, and /ŋ/ have the same sound quality as the Eng. sounds. Simple /n/ occurs only in loanwords. /m̠/ is pronounced in the same way as /m/ but with the tongue pulled back. WOL /l/ is a flap sound more or less like the r in British 'very'. /l/ is pronounced with much less pressure than in the case of /t/. Although /l/ is not a lateral sound in WOL, it corresponds to the lateral /l/ in neighbouring languages including Ulithian. It is for this reason that the symbol /l/ is used to represent the sound.

Observe the following words in which the consonants discussed above are italicised. The examples are given in both phonemic and phonetic transcriptions. The wedge (ˬ) under a symbol indicates voicing.

| | | |
|---|---|---|
| /paapaa/ | [paapa] | *to count* |
| /metai/ | [metai] | *my eyes* |
| /icii/ | [ici] | *one* (from Jp.) |
| /kacitoo/ | [kacito] | *movie* (from Jp.) |
| /yaafi/ | [yaafi̥] | *fire* |
| /silei/ | [silei] | *my mother* |
| /iixe/ | [iixe̥] | *fish* |
| /wexare/ | [wexare̥] | *root* |
| /taaβu/ | [taaβu̥] | *taboo* |
| /taββeeye/ | [taββeeye̥] | *follow him* |
| /šoo/ | [šo] | *copra* |
| /rixi/ | [rixi̥] | *to run* |
| /yaremate/ | [yaremate̥] | *person* |
| /sensei/ | [sensei] | *teacher* (from Jp.) |
| /ŋii/ | [ŋi] | *tooth* |
| /m̀aale/ | [m̀aale̥] | *man* |
| /yaalo/ | [yaalo̥] | *sun* |

There are eleven double consonants as shown below.

| manner of articulation | | | place of articulation | | | | |
|---|---|---|---|---|---|---|---|
| voice | mode | tongue shape | lips | lip-teeth | teeth-gum | hard palate | soft palate |
| voiceless | stop | plain | /pp/ | | /tt/ | /cc/ | /kk/ |
| | | pulled back | /p̈p̈/ | | | | |
| | fricative | plain | | /ff/ | /ss/ | | |
| voiced | nasal | plain | /mm/ | | /nn/ | | /ŋŋ/ |
| | | pulled back | /m̈m̈/ | | | | |

All double consonants and their corresponding simple consonants share the same sound qualities except that double ones are stronger and longer than their simple counterparts. /p̈p̈/ is the only consonant which does not have a simple counterpart /p̈/ but corresponds instead to /β/. It is not only stronger and longer than /β/, but it is also different in that it is a stop sound while /β/ is a fricative. On the other hand, the five simple consonants /β, x, š, r, l/ do not have corresponding double consonants. When /c/, /k/, and /n/ occur in native words, they are always pronounced as double consonants. Observe the following examples in which double consonants are italicised.

| | | |
|---|---|---|
| /ppiye/ | [ppiyę] | *sand* |
| /ttiri/ | [ttiri̥] | *fast* |
| /ccaa/ | [cca] | *blood* |
| /kkaŋi/ | [kkaŋi̥] | *sharp* |
| /p̈p̈eš̈e/ | [ppeš̈ę] | *white* |
| /ffaxe/ | [ffaxę] | *to cough* |
| /ssaaɒu/ | [ssaaɒu̯] | *fishing kit* |
| /m̈m̈ate/ | [mmatę] | *to awake* |
| /ɒunno/ | [ɒunnǫ] | *heart* |
| /ŋŋawe/ | [ŋŋawę] | *bad* |
| /m̈m̈uto/ | [m̈m̈utǫ] | *to vomit* |

A double consonant is slightly weakened if it occurs in word-initial position and is followed by a vowel and a second occurrence of the same double consonant. For instance, /m̈m̈om̈m̈oŋoo/ *to be eating* is pronounced as [m̈om̈m̈oŋo] and /kkekkeɒate/ *to be barking* as [kekkeɒatę].

(2) WOL Vowels

WOL vowel phonemes are either single or double. The six single vowels of WOL may be arranged according to tongue height, tongue frontness or backness, and lip shape, as follows:

| tongue | front | central | | back |
|---|---|---|---|---|
| | (flat lips) | (plain lips) | (round lips) | (round lips) |
| high | /i/ | | /ü/ | /u/ |
| mid | /e/ | | | /o/ |
| low | | /a/ | | |

The phoneme /i/ is similar to a short variety of Eng. 'ee' as in meet, /e/ to Eng. 'e' in bed, /a/ to Eng. 'a' as in park, /ü/ to German ü in müssen *must*, /u/ to the short variety of Eng. 'oo' in pool, and /o/ to Eng. 'o' in short, /e/ is pronounced with slight lip rounding when it occurs before a round vowel, e.g. /iteü/ [iteü] *who?*.

Simple vowels are pronounced as voiceless when they occur in word-final position after a consonant or semivowel, e.g. /maate/ [maatę] *eyes*. The only exception is words of the shape (C)V (optional simple consonant + obligatory vowel) in which case no devoicing occurs, e.g. /me/ [me] *and*, /xo/ [xo] *you*, and /ɒe/ [ɒe] *will*, all of which are function words. Except for /a/, all the simple vowels may occur in word-final position. The voiceless /e/, i.e. [ę], is pronounced with slight lip rounding when it occurs after a round vowel, e.g. /ü̈we/ [ü̈wę] *neck*.

Although native speakers are reluctant to spell word-final voiceless vowels, these vowels are significant in distinguishing words with different meanings. For instance, the two members in each of the following pairs are pronounced differently. Their meanings are different too, indicating that they are distinct words.

| | | | |
|---|---|---|---|
| a. | /peeši/ | [peeš̥i̥] | *hotness* |
| | /peeše/ | [peeš̥e̥] | *lime* |
| b. | /xʊ̊ʊsʊ̊/ | [xʊ̊ʊsʊ̥̊] | *louse* |
| | /xʊ̊ʊse/ | [xʊ̊ʊse̥] | *octopus* |
| c. | /laaṁu/ | [laaṁu̥] | *mosquito* |
| | /laaṁo/ | [laaṁo̥] | *lagoon* |
| d. | /xaccʊ̊/ | [xaccʊ̥̊] | *good* |
| | /xacce/ | [xacce̥] | *fish smell* |
| e. | /metaxi/ | [metaxi̥] | *pain* |
| | /metaxʊ̊/ | [metaxʊ̥̊] | *fear* |
| f. | /laaŋi/ | [laaŋi̥] | *sky* |
| | /laaŋo/ | [laaŋo̥] | *house fly* |

Word-final voiceless vowels are also significant in that they affect the sound quality of the preceding vowel. For instance, the a's and e's preceding a final round voiceless vowel (i.e. [o̥], [ʊ̥̊], [ʊ̥]) are pronounced with the lips quite rounded, as illustrated by the rounded a [ɒ] or rounded e [æ] in each word below:

| | | |
|---|---|---|
| /ṁalo/ | [ṁɒlo̥] | *to hide* |
| /mešaro/ | [mešɒro̥] | *dirt* |
| /ssexʊ̊/ | [ssæxʊ̥̊] | *full* |
| /ɒɒelʊ̊/ | [ɒɒælʊ̥̊] | *dirt* |
| /lewemu/ | [lewæṁu̥] | *your tongue* |

Voiceless word-final vowels are sensitive to the sound environment. If a word follows them closely in the same phrase, they become voiced. In the following expressions, notice the italicised symbols which are voiced vowels.

| | | |
|---|---|---|
| /iṁe kkaa/ | [iṁe kka] | *these houses* |
| /rixi laxo/ | [rix̥*i* laxo̥] | *to run away* |
| /ɒunno xaccʊ̊/ | [ɒunno x̥accʊ̥̊] | *good heart* |
| /xotɔɔto laa/ | [xotɔɔto la] | *that crack* |
| /ɒušo xemase/ | [ɒušo x̥emase̥] | *very stupid* |
| /yalʊ̊sʊ̊lʊ̊ yaape/ | [yalʊ̊sʊ̥̊lʊ̊ yaape̥] | *ghost of Yap* |

There are eight double vowels in WOL which may be arranged as follows:

| tongue | front | central | | back |
|---|---|---|---|---|
| | (flat lips) | (plain lips) | (round lips) | (round lips) |
| high | /ii/ | | /ʊ̊ʊ̊/ | /uu/ |
| mid | /ee/ | | /əə/ | /oo/ |
| low | | /aa/ | | /ɔɔ/ |

A double vowel is about twice as long as a simple vowel. Notice in the above table that there are two double vowels, /əə/ and /ɔɔ/ which do not correspond to any simple vowel phonemes. These two double vowels were historically each a sequence of two different simple vowels, but in modern WOL, they are simply long vowels. /əə/ is similar to French oeu in coeur *heart*, pronounced with the lips rounded. It is the sound made by simultaneous pronunciation of /ee/ and /oo/. /ɔɔ/ is similar to Eng. 'aw' in law. All the other double

vowels are simply lengthened varieties of the corresponding simple ones. While simple vowels in word-final position are devoiced after a consonant or semivowel, double vowels are slightly shortened, with voicing retained, in the same position. Thus, in the following examples, word-final (short) voiced vowels which occur after a consonant or semivowel are actually 'allophones' of double vowels.

| | | |
|---|---|---|
| /iiyaa/ | [iiya] | *where?* |
| /leelixuu/ | [leelixu] | *a taro sp.* |
| /yaai/ | [yaai] | *mine* |
| /lʉʉlʉʉ/ | [lʉʉlʉ] | *to chew* |
| /ləələə/ | [ləələ] | *to fill with liquid* |
| /tuuloŋo/ | [tuuloŋg] | *to dive in* |
| /ƥooti/ | [ƥotį] | *nose* |
| /lɔɔlɔɔ/ | [lɔɔlɔ] | *wavy* |

(3) WOL semivowels

There are two semivowels, /w/ and /y/. They are glides in that the tongue starts at a position slightly higher than for /u/ and /i/, respectively, and approaches the position for the vowel that immediately follows. /w/ and /y/ are phonemes in that they serve to differentiate meanings, as shown in the following minimal pairs.

| | | |
|---|---|---|
| /waai/ | [waai] | *my canoe* |
| /yaai/ | [yaai] | *mine* |
| /woo/ | [wo] | *bunch* |
| /yoo/ | [yo] | *to sneak* |

(4) Phoneme alternations

While there are non-significant (allophonic) variations in sounds, as between [i] and [į], there are also alternations between phonemes. As in allophonic variations, phoneme alternations are caused in many cases by the sound environment in which they occur. In what follows, let us briefly observe such phoneme alternations in WOL.

As already stated, the five simple consonants /ƥ/, /x/, /š/, /r/, and /l/ do not have corresponding double consonants. If, for some grammatical purpose such as the formation of the progressive form of a verb, these consonants are doubled, a doubled /ƥ/ becomes /ƥ̇ƥ̇/, doubled /x/ becomes /kk/, doubled /s/ and /r/ both become /cc/, and doubled /l/ becomes /nn/, as illustrated below.

| | | |
|---|---|---|
| /ƥuxaa/ | [ƥuxa] | *boil it*: /ƥ̇ƥ̇uƥ̇ƥ̇uxaa/ |
| | [ƥ̇uƥ̇ƥ̇uxa] | *be boiling it* |
| /xaŋii/ | [xaŋi] | *eat it*: /kkekkaŋiiye/ |
| | [kekkaŋiiyę] | *be eating it* |
| /šaalʉ/ | [šaalʉ̨] | *water*: /ccecclʉʉwe/ |
| | [ceccalʉʉwę] | *fill it with water* |
| /raaŋe/ | [raaŋę] | *yellow powder*: /ccecc aŋe/ |
| | [ceccaŋę] | *apply yellow powder* |
| /lʉtʉ/ | [lʉtʉ̨] | *to jump*: /nnʉtʉ/ |
| | [nnʉtʉ̨] | *to be jumping* |

In contrast to these grammatically conditioned alternations, there are consonant alternations which are phonologically conditioned. For instance, /mannʉ/ [mannʉ̨] *bird of* results from /malʉ/ *bird* + /li/ *of* with the deletion of the vowel /ʉ/ before a morpheme boundary and between /l/'s. With the deletion of /ʉ/, two /l/'s are automatically changed to /nn/. Also it is clear that /ʉ/

causes the following vowel /i/ to be changed to /ů/ before it is deleted, hence /mannů/. Thus, we can say that /malů-li/ is the *base form* from which /mannů/ is derived. Base forms will be preceded by an asterisk in what follows e.g. *malů-li *bird of*. Observe the following examples.

*xala-la /xanne/ [xannę] *his food*
*tapo-li-šůů-lapa /tapoccůůlape/ [tapoccůůlapę] *tip of the backbone*

Notice in the second example that, with the deletion of i, the consonants l and s are collapsed to /cc/. In the same way, l + r = /cc/, l + t = /tt/, and l + s = /ss/. Vowel deletions will be discussed shortly under vowel alternations.

Rather extensive phoneme alternations may be observed in vowels. For instance, the word which means 'name' is phonemically represented in three different ways, /iite/, /ite/, and /ita/. The form /iite/ shows up when the word is pronounced without any modifying element following. The form /ite/ appears in three different environments: (i) before a modifying word, as in /ite xacců/ *good name*; (ii) before a suffix whose initial vowel is high, as in /itei/ *my name*; and (iii) before a suffix whose initial vowel is /a/, as in /itemami/ *our* (excl.) *name*. The form /ita/ appears in all other environment not mentioned above, as in /itale/ *his name*. Since the different realisations of the word 'name' are due to sound environments such as neighbouring vowels, the word boundary, and the presence or absence of a modifying word, we can easily imagine that there is a *base form* for each word which is free from environmental influence, and that the different forms associated with each word are predictable from the base form in terms of their sound environments. For instance, suppose we set up /ita/ as the base form for 'name'. Then /iite/ and /ite/ are derivable from the base form *ita by means of a few general phonological rules which will be given shortly. We cannot adopt /iite/, /ite/, or any other form as the base form, because none of them can explain the derivation of the other forms.

Technical details aside, the following simple procedures may be used to obtain base form vowels of WOL words.

a. If there is an alternation between /a/ and other vowels, choose /a/ as the base form vowel. For instance, between /e/ and /a/ in /itei/ *my name* and /itale/ *his name*, /a/ should be the base form vowel, hence *ita *name*.

b. If a long first vowel of an independently pronounced word, e.g. /iite/ *name*, alternates with the corresponding short vowel contained in a larger word, e.g. /itale/ *his name*, the short vowel should be taken for the base form, e.g. *ita *name*. If a long vowel does not have a short alternate vowel, the long one is the base form vowel. For instance, in /ƥooti/ *nose* and /ƥootile/ *his nose*, the first vowel is consistently long, hence *ƥooti *nose*.

c. While a voiced final vowel of a content word is represented as long in both base and phonemic form, e.g. *waa /waa/ *canoe*, a voiced final vowel of a function word of (C)V type is represented as short, e.g. *i /i/ *I*, *xo /xo/ *you*, *ye /ye/ *he, she, it*, and *ŋe /ŋe/ *and*. This is because a final vowel of a content word is slightly longer than that of a function word even when pronounced in isolation, and the former is definitely lengthened when followed by some other element, e.g. [waalę] *his canoe*, but the latter remains always short, whatever environment it may occur in.

d. If there is no phoneme alternation, base form vowels should be based on the pronunciation, e.g. *yase /yaase/ [yasę] *liver*, *yalůsů /yalůsů/ [yalůsų] *ghost*, and *taai /taai/ [taai] *no longer*.

e. Many words and parts of words do not allow any element to be attached after them, preventing us from observing the alternation of final voiceless vowels. The general rule in such cases is simply to set up the voiced counterparts of the voiceless vowels for the base forms. Thus, for instance, the base forms of /šaxʊ̈/ [šaxʊ̥] *just*, /xaaŋʊ̈/ [xaaŋʊ̥] *I*, /m̊u/ [m̊ʊ̥] *your*, and /li/ [liᵢ] *of* are respectively *šaxʊ̈, *xaaŋʊ̈, *m̊u, and *li. One exception to this rule is that the base form vowel of a final /e/ [ɐ̥] is a, e.g. *la /le/ [lɐ̥] *his*. This exception is based not only on historical and comparative evidence but on the fact that all other final a's are realised as /e/ anyway.

Once base forms for all WOL morphemes are determined, the following major phonological rules, among others, will derive the phonemic vowels from base form vowels. None of the following rules applies to base forms of (C)V type.

(a) A simple word-final a becomes /o/ after o, ɔɔ, or u + consonant.

Examples: *punna /punno/ [punnɔ̥] *heart*
*punna xaccʊ̈ /punno xaccʊ̈/ [punno xaccʊ̈] *good heart*

(b) In all the other environments, a simple word-final a becomes /e/.

Examples: *mata /maate/ [maatɐ̥] *eye*
*mata xaccʊ̈ /mate xaccʊ̈/ [mate xaccʊ̥] *good eye*

(c) A simple word-final i becomes /ʊ̈/ after ʊ̈, and /u/ after u.

Examples: *yalʊ̈sʊ̈-li /yalʊ̈sʊ̈lʊ̈/ [yalʊ̈sʊ̈lʊ̥] *ghost of*
*meŋaaxu-li Tonii /meŋaaxulu tonii/ [meŋaaxulu toni] *Tony's clothes*

(d) The first simple vowel of a base form becomes doubled (or lengthened) when there is no modifier following and when both of the following conditions are met:

i. The base form must consist of only two simple vowels, with one or two simple consonants or semivowels. Thus, the first vowels of *punna *heart*, *falʊ̈wa *island*, *pee *divination*, do not undergo doubling.

ii. The base form must be used as a noun. For instance, the first vowels of the verb *ʊ̈lʊ̈ *to drink* and the adverb *šaxʊ̈ *just* do not undergo doubling.

Examples: *fau /faaʊ̈/ [faaʊ̈] *stone*
*ita /iite/ [iitɐ̥] *name*
*ʊ̈wa /ʊ̈ʊ̈we/ [ʊ̈ʊ̈wɐ̥] *neck*
*mata /maate/ [maatɐ̥] *eye*

(e) A simple a becomes /e/ between high vowels. This rule applies across word boundaries as long as the word boundary is not accompanied by a breath pause.

Examples: *itai /itei/ [itei] *my name*
*raxiraxi /raxirexi/ [raχirexiᵢ] *to line up*
*malʊ̈ kkaila /malʊ̈ kkeile/ [malʊ̈kkeilɐ̥] *strong man*
*xaaʊ̈ ŋalii /xaaʊ̈ ŋelii/ [xaaʊ̈ŋeliᵢ] *to say to him*
*i maŋimaŋi /i meŋimeŋi/ [imeŋimeŋiᵢ] *I think*

(f) When the simple vowel a occurs before a low vowel (i.e. /a/ or / /), it becomes /e/. This rule operates from right to left (Sohn 1971).

Examples: *mata-i /metai/ [metai] *my eyes*
*marama /merame/ [merameˌ] *moon*
*xamɔɔa /xemɔɔwe/ [xemɔɔweˌ] *erase it*

(g) Before a morpheme (i.e. minimal meaningful unit) boundary, a simple vowel drops obligatorily between two identical consonants and optionally between l and a teeth-gum or palatal consonant.

Examples: *malu̇-li /mannu̇/ [mannu̇ˌ] *bird of*
*sila-la /sinne/ [sinneˌ] *his mother*
*lima-malu̇ /limmelu̇/ [limmelu̇ˌ] *five animate objects*
*tama-li-tama-i /temattemai/ [temattemai] *my grandfather*
*yau̇-li-rale /yau̇ccale/ [yau̇ccaleˌ] *weather*
*tapo-li-šu̇u̇ /tapoccu̇u̇/ [tapoccu̇ˌ] *bone tip*

(h) Before the third person plural possessive suffix *ra /re/, a simple vowel is always lengthened and, moreover, if the vowel is a preceded by a high vowel, it goes to /ee/, not /aa/.

Examples: *yawa-ra /yewaare/ [yewaareˌ] *their mouths*
*upa-ra /upeere/ [upeereˌ] *their chests*
*lewe-ra /leweere/ [leweereˌ] *their tongues*
*sila-ra /sileere/ [sileereˌ] *their mother*

Phoneme alternation is also observable in semivowels. There are certain morphemes which in their initial position have /w/ in one environment, /y/ in another, and zero in the third, without any change in meaning. This alternation is perfectly predictable, given the base forms of the meaning units and the environments in which they occur. That is, /w/ appears if the preceding vowel is round, /y/ appears if the preceding vowel is unrounded, and nothing appears if the preceding vowel is the same simple vowel as the following one. For instance, the base form of the unit meaning *me* is *ai and that of the unit meaning *it* is *a. The base form *ai *me* corresponds to the four alternant forms /wei/, /wai/, /yei/, and /yai/. The base form *a *it* corresponds to /we/, /ye/, and /a/. The /w/-/y/-zero alternation is due to the environments stated just above. The alternation /e/-/a/ is due to the environments mentioned with regard to vowel alternations. Examples follow.

*xamasu̇ru̇u̇-ai /xemasu̇ru̇u̇wei/ [xemasu̇ru̇u̇wei] *make me sleep*
*xa̰mo̰o-ai /xa̰mo̰oowai/ [xa̰mo̰oowai] *feed me*
*weri-ai /weriyei/ [weriyei] *see me*
*xu̇laa-ai /xu̇laayai/ [xu̇laayai] *know me*
*xasu̇u̇we/ [xasu̇u̇weˌ] *build it*
*liia-a /liiye/ [liiyeˌ] *kill it*
*nu̇u̇wa-a /nu̇u̇waa/ [nu̇u̇wa] *wash it*

## 3. THE PROPOSED WOL ALPHABET

An orthography has two aspects, paradigmatic (systematic) and syntagmatic (combinatorial). The two aspects correspond respectively to sound systems and sound combinations. The paradigmatic aspect, to be discussed in this section, is associated with a set of graphic symbols or letters, whereas the syntagmatic aspect, which will be taken up in Section 4, is associated with spellings or spelling conventions.

(1) The following alphabetic and other relevant symbols are proposed.

Consonants

| simple | alphabetic symbols | double | alphabetic symbols |
|---|---|---|---|
| /p/ | p | /pp/ | pp |
| /t/ | t | /tt/ | tt |
| /c/ | c | /cc/ | cc |
| /k/ | k | /kk/ | kk |
| /f/ | f | /ṗṗ/ | bb |
| /s/ | s | /ff/ | ff |
| /x/ | h | /ss/ | ss |
| /ṗ/ | b | /mm/ | mm |
| /š/ | sh | /nn/ | nn |
| /r/ | r | /ŋŋ/ | nng |
| /m/ | m | /ṁṁ/ | mmw |
| /n/ | n | | |
| /ŋ/ | ng | | |
| /ṁ/ | mw | | |
| /l/ | l | | |

Vowels

| simple | alphabetic symbols | double | alphabetic symbols |
|---|---|---|---|
| /i/ | i | /ii/ | ii |
| /e/ | e | /ee/ | ee |
| /a/ | a | /aa/ | aa |
| /u̇/ | iu | /u̇u̇/ | iiu |
| /u/ | u | /əə/ | eo |
| /o/ | o | /uu/ | uu |
| | | /oo/ | oo |
| | | /ɔɔ/ | ao |

Semivowels

| | |
|---|---|
| /w/ | w |
| /y/ | y |

Other symbols

. period
, comma
? question mark
! exclamation mark
() parentheses
" " or ' ' quotation marks
- hyphen

Arabic numerals are used to indicate numbers.
Mathematical symbols like +, -, and ± are used.
Capital letters are used for the first letters in sentences and proper nouns.

(2) All the symbols given above are borrowed from the alphabetic and other symbols used in writing Eng. By borrowing them from Eng., time and labour will

be saved in learning to read and write, not only due to our familiarity with them but also due to their simplicity in comparison with other alphabets existing in the present-day world. Furthermore, adoption of the letters from the Eng. alphabet allows us to use machines for the graphic representation of Eng., including typewriters.

The letters of the Eng. alphabet, however, do not provide all the symbols required by the significant WOL sounds. In order to get around such a situation, it is a common practice to use two letters to indicate other sounds, or to introduce some kind of superscript such as (') , (") , and (¯). The former we call digraphs and the latter, diacritic marks. In general, learning to discriminate digraphs (consisting of two existing symbols) is easier than learning to discriminate diacritics which are entirely new symbols, since digraphs are mnemonically much better. Moreover, digraphs are easier to type than diacritics. One serious drawback of digraphs, however, is that ambiguities may occur as in Eng. *ph* (h*ap*hazard) vs. *ph*oto). If such ambiguities are suppressed, digraphs are preferred. In WOL, it is proposed not to include diacritic marks. Digraphs are allowed to the extent that they do not bring in ambiguity. The digraphs are iu, iiu, eo, ao, sh, mw, mmw, ng, and nng. The reasons for proposing these symbols are summarised below. Throughout this section, asterisked morphemes or words represent spellings.

The digraph iu for /u̇/ (as in *faiu /faau̇/ *stone*) is introduced to distinguish the vowel from /i/ and /u/. The fact that the phonetic quality of /u̇/ is ambivalent between those of /i/ and /u/ justifies the composition of the digraph. An alternative might be to set up u for /u̇/ and wu for /u/, since it is more economical in terms of time, space, and labour in reading and writing to assign complex symbols to those sounds like /u/ which occur rarely, and simple symbols to frequently appearing sounds like /u̇/. Another supporting reason for this alternative might be that the original /u/ of the Oceanic languages corresponds to WOL /u̇/, whereas WOL /u/ is a later development. This alternative is not adopted, however, because then *mwu would be ambiguous between *m-wu /mu/ and *mw-u /mu̇/. Actually, /mu/ does not occur in WOL while /mu̇/ does (e.g. /mu̇ccu̇/ *to complete*), which would allow us to establish a convention to the effect that, in such sequences, w always goes with u. Even in that case, writing /mu/ *your* as *mwwu would look very inelegant.

Both eo and ao (as in *leoleo /ləələ/ *fill with liquid* and *laolao /lɔɔlɔɔ/ *wavy*) represent single long vowels. There is no short counterpart of either. The use of two vowel symbols is justified by the fact that they are long. Selection of the composite symbols a and o for ao and e and o for eo is motivated by the fact that the sound represented by eo is pronounced with the tongue in a position intermediate between that for e and that for o, while the sound represented by ao is pronounced with the tongue low as for a, but with the lips rounded as for o. Eo and ao, wherever they appear, are each pronounced in only one way, and therefore there is no ambiguity. The WOL sound pattern does not permit sequences e and o, or a and o, without an intervening consonant or semivowel.

An alternative might be to reverse the letters as in oa instead of ao and oe instead of eo. This has to be decided upon during the upcoming orthography meeting.

The choice of the symbol sh is motivated by tradition as well as by similarity to the sound sh in the Eng. word 'shoes'. However, there are other alternatives that are equally plausible. It is strongly recommended that we use the same symbols for all the dialects of the Outer Islands so far as the sounds represented by sh and cc (or c) are concerned, because the interisland pronunciation differences are quite regular and predictable as shown below in:

| WOL | ULI | Satawal | Alternative spellings | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | I | II | III | IV | V | VI | VII |
| [š] | [c] | [ṛ] | sh | sh | sh | sh | c | ch | ch |
| [cc] | [cc] | [cc] | cc | cc | c | ch | cc | cch | cc |
| ([c]) | | | (c) | | | | | | |

As already mentioned, WOL /c/ occurs only in loanwords. The decision has to be made at the orthography meeting.

Mw which is pronounced with the tongue back approaching the velum is distinguished from plain m. Tradition favours the digraph, too.

As regards ng, our alphabet follows tradition, because the alternative is commonly understood as indicating a non-nasal sound and sometimes people use g for the sound represented by h (as in *laho *to go*).

Double sounds related to the above digraphs are represented by doubling the first symbol, as in iiu, mmw, and nng.

Besides the above digraphs, selection of simple symbols has also been given careful attention, particularly the following:

h. According to tradition, both h and g have been used for the same sound. Sometimes, one person may use both indiscriminately. The reason for proposing h is simply that the sound is fricative and basically voiceless, although it becomes voiced between two voiced vowels.

b and bb. The sound represented by b /ḅ/ is basically a voiceless fricative, but it becomes voiced between voiced vowels. While this sound is a fricative, its long counterpart, bb /ḍḍ/ is a stop. There are several other alternative spellings to consider:

| | Short | Long |
|---|---|---|
| 1. | b | pw |
| 2. | pw | ppw |
| 3. | fw | ffw |
| 4. | bw | bbw |

Choice of b and bb is motivated by simplicity.

l. Although the sound is more or less like a flap [ṛ] as already mentioned, we adopted the symbol l in view of the fact that the sound corresponds to [l] in the neighbouring languages and there is a separate sound, the resonant [r], to which the symbol r is more appropriately assigned.

## 4. SPELLING CONVENTIONS

(1) Basic principles.

The next consideration is how to spell WOL words and sentences by means of the abovementioned alphabetic letters. In the first place, we have to decide in what manner sounds and letters are to be matched in actual words. There may be several alternatives.

One is to write down words in letters corresponding to their actual pronunciation. Since, for instance, the word for 'name' is pronounced as [iitę], this phonetic transcription would be used as the spelling for the word. This type of writing may be called the phonetic spelling. However, phonetic spelling would require the writer spelling in this way to pay an inordinate amount of attention to purely automatic allophonic variations which may be predicted from the sound environment in which they occur.

A second alternative is to spell words in letters corresponding to their phonemic representation. Since, for instance, the word for 'name' is represented phonemically as /iite/, the sequence *iite would be used as the spelling. In the same way, *my name* is *itei, *his name* *itale, and *our* (excl.) *names* *itemami. This type of spelling convention may be called phonemic spelling.

A third alternative is to write down words in letters corresponding to their base form representation. For instance, the word for 'name' would be represented as *ita. In the same way, *my name* would be *itai, *his name* *itala and *our* (excl.) *names* *itamami. As we have seen, base forms of words are set up by reducing any sounds caused by environments to their source sounds, that is, to those sounds which would be pronounced if no sound environment were present. This type of writing may be called base form spelling.

Other alternatives include combinations of the above three types of spelling, which will not be discussed here. Let us compare the phonemic and base form spellings to observe their merits and shortcomings.

As we saw in the examples just above, phonemic spelling is more faithful to actual pronunciation than base form spelling is. On the other hand, base form spelling is very neat in that, in general, one morpheme may be spelled in only one way, which is not the case with phonemic spelling. One of the most serious drawbacks of the latter system is that there must be two or three or more spellings for a great many WOL morphemes. Our example *name* must be spelled as three different ways, *iite, *ite, and *ita in spite of the fact that the differences are entirely predictable from the sound environment. The causative prefix must also have two different spellings, he- and ha- although the first variant may be derived from the second by the a-raising that is conditioned by a low vowel in the following syllable.

In general, a writing system need not reflect what native speakers can predict by general phonological rules (c.f. Lester 1974:141). This is tantamount to saying that a writing system should be a base form spelling, in which one morpheme is represented only by one spelling. Even Eng., in spite of so many irregular correspondences between sound and letter, has adopted an extensive degree of base form spelling (e.g. Chomsky and Halle 1968:48). Compare the two types of spelling below and notice how the actual practice follows the base forms.

| rough pronunciation | phonemic spelling | base form spelling |
|---|---|---|
| [pæts] | pats | *pats* |
| [boyz] | boyz | *boys* |
| [æksɨz] | axɨz | *axes* |
| [peintid] | peint*i*d | *painted* |
| [workt] | workt | *worked* |
| [pleid] | plei*d* | *played* |
| [politik] | politi*k* | *politic* |
| [politišən] | politi*sh*ian | *politician* |

The one-morpheme one-spelling principle of the base form spelling has the following advantages among others.

(a) Reading and writing for meaning will be greatly facilitated, since one-morpheme one-shape contributes to quick word-recognition as well as to quick memorisation of spellings. This applies both to native speakers and to language learners. Observe the 'plural' suffix -s in the above Eng. examples. It has three different pronunciations [s], [z], and [ɨz], yet all of them are spelled the same way, i.e. s. As it is spelled in only one way, it allows a more rapid translation from spelling to meaning and vice versa than would spelling it in three different ways, say, s, z, and ɨz. One might say that this kind of one-morpheme one-shape spelling is a hindrance to those who are in the initial stages of learning to read. However, there is no evidence that such a system is a barrier to learning to read, even in the initial stages. Once a few simple general rules of pronunciation are observed, together with instructions as to how each letter is to be pronounced, it will be easy for the reader to pronounce each word correctly. Moreover, the ultimate goal of reading is not rapid pronunciation but quick recognition of meanings of words and sentences, to which one-morpheme one-spelling definitely contributes, since such a system will maintain a close correspondence between morphemes and their orthographic representations.

(b) The one-morpheme one-spelling principle will contribute to the simplicity of dictionaries. Instead of listing different forms of one word as two, three, or four separate entries, we need list only one form, and the rules of predictable pronunciations can be given in a few lines at the beginning of the dictionary. It would be pointless and clumsy for the orthography to indicate such predictable forms and include them in the dictionary as separate entries.

(c) Furthermore, in many cases the base form spelling allows different dialects to have the same or very similar spellings in spite of considerable differences in pronunciation. Predictable sound changes are in many cases limited to certain dialects and not to others. If such predictable sounds are not represented in spellings, the same spellings can be used for many dialectally different words. For example, observe the difference in phonemic representation in the following words:

| ULI | WOL | | Spelling |
|---|---|---|---|
| /maram/ | /merame/ | *moon* | marama |
| /maramal/ | /maremali/ | *moon of* | maramali |

If we spell /e/ in WOL dialect as a in view of the predictability of /e/ from a, the same spelling can be used for both ULI and WOL dialects in spite of the difference in pronunciation. One important consequence of this is that speakers of one dialect can read texts written in the other dialect easily because they recognise many words which are spelled the same way in spite of different pronunciations. This situation exists in many other languages, including Eng. and Chinese. For example, Eng. is spoken in many different parts of the world, and speakers from one place, say Texas, may not at first understand easily the Eng. spoken by a farmer in New Zealand. In spite of that, written Eng. is easily understood by everyone. What this implies is that if the Administration of Yap District were to make written announcements, say, in the ULI dialect, then the Woleaian and Satawalese people would have no difficulty in reading them because of the many words spelled in the same way as in their own dialect.

Word boundaries will be marked by spaces. What constitutes a word for spelling purposes is determined by the sound patterns and grammatical structures of WOL (see Sohn 1975 for details). However, traditional practices will also have to be taken into account so far as options are allowed. In principle, each content or function word with or without affixes is considered a spelling unit, as shown below.

*Paangali yaremata nge re ccapara be re laho teramiya.
/paaŋali yaremate ŋe re ccepare ƥe re laxo teramiye/
*Everybody believes that he will go to heaven.*

*Semaliu maliumwashoho ye liia me lanni ruumwu wee yaala.
/semalu̯ malu̯mašoxo ye liiye me lanni ruu̯mu wee yaale/
*A thief killed him in his room.*

If two or more geographical or social dialects have different pronunciations for the same meaning, two cases may arise. First, the pronunciations may not be related at all, as in the use of WOL *ranga and Faraulep taiho for *turmeric*. Second, the pronunciations are so related that the general phonological rules we mentioned are able to reduce them to the same spelling, then they are automatically spelled the same way as in the case of ULI /maram/ and WOL /merame/ both as *marama *moon*. If they cannot be reduced to the same spelling, they have to be spelled differently:

Eastern WOL *peleehisi *blanket*
Western WOL *pelaihiti "

Which dialect of the Outer Islands is to be used for 'standard' or 'official' purposes should be decided by the speakers of the dialects taking into consideration political structure, geographical locations, population, etc.

(2) Examples.

In view of the advantages discussed above, I propose that the WOL spelling conventions follow, in principle, the base form spelling. Relegating some other spelling-related conventions to the appendix, let us observe some examples of base form spellings below.

| pronunciation | glosses | phonemic representation | proposed spelling |
|---|---|---|---|
| [u̯u̯we̯] | *neck* | /u̯u̯we/ | iuwa |
| [u̯wei] | *my neck* | /u̯wei/ | iuwai |
| [u̯wale̯] | *his neck* | /u̯wale/ | iuwala |
| [walu̯welu̯] | *plant* | /walu̯welu̯/ | waliuwaliu |
| [iteilaxo̯] | *I don't go.* | /i tei laxo/ | I tai laho. |
| [yetailaxo̯] | *He doesn't go.* | /ye tai laxo/ | Ye tai laho. |
| [merame̯] | *moon* | /merame/ | marama |
| [maremali̯] | *moon of* | /maremali/ | maramali |
| [tepaŋi] | *help him* | /tepaŋii/ | tapangii |
| [xetapetape̯] | *to help* | /xetapetape/ | hatapatapa |
| [xatə] | *make (him) climb* | /xatəə/ | hateo |
| [xemɔowe̯] | *erase it* | /xemɔowe/ | hamaoa |
| [ƥunno̯] | *heart* | /ƥunno/ | bunna |
| [ƥunnale̯] | *his heart* | /ƥunnale/ | bunnala |
| [kku̯] | *fingernail* | /kku̯/ | kkiu |
| [kku̯u̯] | *my fingernail* | /kku̯u̯/ | kkiui |
| [kkutu] | *my spit* | /kkutuu/ | kkutui |
| [meŋaaxulu̯] | *clothes of* | /meŋaaxulu/ | mengaahuli |

| pronunciation | glosses | phonemic representation | proposed spelling |
|---|---|---|---|
| [peše] | *foot* | /pešee/ | peshee |
| [i] | *I* | /i/ | i |
| [si] | *we* (incl.) | /si/ | si |
| [xaaŋų̊] | *I* (ind. form) | /xaaŋų̊/ | haangiu |
| [šę̊] (suffix) | *our* (incl.) | /še/ | sha |
| [rę̊] (suffix) | *their* | /re/ | ra |
| [yai] (suffix) | *me* | /yai/ | ai |
| [yei] (suffix) | | /yei/ | |
| [wai] (suffix) | | /wai/ | |
| [wei] (suffix) | | /wei/ | |
| [yę̊] (suffix) | *him, her, it* | /ye/ | a |
| [wę̊] (suffix) | | /we/ | |
| [ ] (suffix) (e.g. [xų̊la_] *know him*) | | /a/ (e.g. /xų̊laa/) | |
| [tapeeyę̊] | *follow him* | /tapeeye/ | tabeea |
| [yaăwę̊] | *mouth* | /yaawe/ | yawa |
| [yewaarę̊] | *their mouths* | /yewaare/ | yawara |
| [siilę̊] | *mother* | /siile/ | sila |
| [silalę̊] | *his mother* | /silale/ | silala |
| [sileerę̊] | *their mother* | /sileere/ | silara |
| [limmelų̊] | five animate objects | /limmelų̊/ | limamaliu |
| [imettamolų̊] | *the chief's house* | /imettamolų̊/ | imwalitamwoliu |
| [wessilei] | *my mothers* | /wessilei/ | welisilai |
| [weccapį̊] | *ancestors* | /weccapi/ | welishapi |
| [yaų̊ccalę̊] | *weather* | /yaų̊ccale/ | yaiulirala |

# APPENDIX

Following the general recommendations of the Yap Outer Islands Orthography Committee (Kuroiwa 1973), I adopted the following alphabet and spelling conventions for a reference grammar of Woleaian (Sohn 1975) and a Woleaian-English dictionary (Sohn and Tawerilmang 1976). For related discussions, see Sohn (1975:Chapter 2).

(1) The WOL Alphabet

Consonants

| simple | alphabetic symbols | double | alphabetic symbols |
|---|---|---|---|
| /p/ | p | /pp/ | pp |
| /t/ | t | /tt/ | tt |
| /c/ | ch | /cc/ | ch |
| /k/ | k | /kk/ | k |
| /f/ | f | /p̧p̧/ | bb |
| /s/ | s | /ff/ | ff |
| /x/ | g | /ss/ | ss |
| /p̧/ | b | /mm/ | mm |
| /š/ | sh | /nn/ | n |
| /r/ | r | /ŋŋ/ | nng |
| /m/ | m | /ṁṁ/ | mmw |
| /n/ | n | | |
| /ŋ/ | ng | | |
| /ṁ/ | mw | | |
| /l/ | l | | |

Vowels

| | | | |
|---|---|---|---|
| /i/ | i | /ii/ | ii |
| /e/ | e | /ee/ | ee |
| /a/ | a | /aa/ | aa |
| /u̇/ | iu | /u̇u̇/ | iu |
| | | /əə/ | eo |
| /u/ | u | /uu/ | uu |
| /o/ | o | /oo/ | oo |
| | | /ɔɔ/ | oa |

Semivowels

| | |
|---|---|
| /w/ | w |
| /y/ | y |

(2) Spelling Conventions.

Almost all native speakers seem to prefer tradition and convenience to linguistic simplicity and clarity, at least at this stage of linguistic sophistication. This is quite understandable when we take into account the popular notion that writing systems are only for those who know the language. Some major conventions are as follows.

a. Final simple vowels which become voiceless in independent forms are not spelled. Thus, for instance, the following three words are to be spelled identically in spite of the differences in pronunciation.

| | | |
|---|---|---|
| [xaccu̥] | *good* | gach |
| [xaccọ] | *tickle* | gach |
| [xaccẹ] | *fish smell* | gach |

Moreover, in base form spelling, the third person singular object suffix would be represented as a, as in *giulaa [xu̇la] *know it*, *liia [liiyẹ] *kill it*, and *gasiiua [xasu̇u̇wẹ] *erect it*. In the spelling adopted, these words are represented as *giula, *liiy, and *gasiuw, respectively, thus entirely obscuring the formal identity of the grammatical morpheme.

b. The vowel lengthening that occurs in nouns of certain sound combinations is disregarded in spelling, as in [iitẹ] /iite/ *name* *it.

c. Except for (a) and (b) above, all sounds are, in principle, spelled as they are pronounced, as in [itei] *my name* *itei, [italẹ] *his name* *ital, [peše] *foot* *peshe, and [me] *and* *me. However, if a word consists only of a double consonant followed by a voiceless vowel, that vowel is spelled, as in [ttị] *to close* *tti. Thus, spellings of [ttị] and [tti] *tea* *tti are not distinguished.

d. The unit of spelling is the word, including compounds. Each word is preceded and followed by a space. If two words, one of which is a modifier, are phonologically close-knit and one cannot occur without the other, like a numeral and a classifier ([semalu̥] 'one animate object' *semal), they form a compound.

(3) Examples.

The examples given in Section 4 (2) are respelled as follows according to the conventions adopted.

*iuw *neck*, *iuwe; *my neck*, *iuwal *his neck*, *waliuwel *plant*, *I tei lag *I don't go*, *Ye tai lag *He doesn't go*, *meram *moon*, *maremal *moon of*, *tepangi *help him*, *getapetap *to help*, *gateo *make (him) climb*, *gemoaw *erase it*, *bun *heart*, *bunal *his heart*, *kiu *finger-nail*, *kiu *my finger-nail*, *kutu *my spit*, *mengaagul *clothes of*, *peshe *foot*, *i *I*, *si *we* (incl.), *gaang *I* (ind. form), *sh *our* (incl.), *r *their*, *yai, *yei, *wai, *wei *me*, *y, *w, *zero *him, her, it*, *tabeey *follow him*, *yaw *mouth*, *yewaar *their mouths*, *sil *mother*, *silal *his mother*, *sileer *their mother*, *limmel 'five anmiate objects', *imwettamwol *the chief's house*, *wessilei *my mothers*, *wechap *ancestors*, *yaiuchal *weather*.

Since the spelling conventions we have adopted are not those of the base form spelling, we have included a base form after each headword in the Woleaian-English dictionary (Sohn and Tawerilmang 1976).

## NOTES

1. This paper is a revised version of the one originally prepared for the Yap Outer Islands Orthography Committee which met in Colonia, Yap, from December 26, 1972 through January 5, 1973. The participants and members of the Committee were:

| | | |
|---|---|---|
| Committee Members: | Henry Ibwol, Ibemai | (Satawal) |
| | Louis Sugmay | (Lamotrek) |
| | Marcus Tewairal | (Faraulep) |
| | Thomas Yanesmal | (Ifaluk) |
| | John Haglelgam | (Eauripik) |
| | Louis Roholesam | (Fais) |
| | Eppolitus Rugwe | (Ulithi) |
| | Anthony Tawerilmang | (Woleai) |
| PALI* Consultants: | Anthony Tawerilmang | (Woleai) |
| | Ramon Peyal | (Ulithi) |
| | Ken Kuroiwa | |

(*Pacific and Asian Linguistics Institute of the University of Hawaii)

## BIBLIOGRAPHY

CHOMSKY, Noam and Morris HALLE

1968 *The sound pattern of English.* New York: Harper & Row.

KUROIWA, Kenneth

1973 Report on the Yap Outer Islands Orthography Committee Meeting. Mimeo. University of Hawaii.

LESTER, Mark

1974 Bilingual education in the United States, the Pacific and Southeast Asia. *Topics in Culture Learning* 2:137-146. (Reprinted in R.W. Brislin, ed. *Culture learning,* 79-88. Honolulu: East-West Center.)

RIESENBERG, Saul H. and Shigeru KANESHIRO

1960 A Caroline Islands script. *Bureau of American Ethnology Bulletin* 173:269-333. Washington, D.C.: Smithsonian Institute.

SMITH, Alfred G.

1951 Gamwoelhaelhi ishilh Weleeya (Guide to Woleai spelling). Mimeo. Department of Education, Office of the High Commissioner of the Trust Territory of the Pacific Islands.

SOHN, Ho-min

1971 *a*-raising in Woleaian. *Working Papers in Linguistics, University of Hawaii* 3/8:15-35.

1975 *Woleaian reference grammar*. Honolulu: University Press of Hawaii.

SOHN, Ho-min and Anthony F. TAWERILMANG

1976 *Woleaian-English dictionary*. Honolulu: University Press of Hawaii.

Sohn, H-M. "An orthographic design for Woleaian". In Bender, B.W. editor, *Studies in Micronesian linguistics*. **C-80:215-236**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.215*

# SELECTING AN ORTHOGRAPHY FOR SAIPAN CAROLINIAN

Frederick H. Jackson

## 1. INTRODUCTION[1]

One of the most important practical tasks which may be asked of a linguist is to assist in the development of an orthography for a language for which no systematic writing system has previously existed. There is a great responsibility inherent in attempting such a task, for the success of a proposed orthography (or lack thereof) may in large part determine administrative and educational policy for the speakers of the language and even the speakers' own attitudes towards the inherent worth of their language.

Although published discussion over the past fifteen years or so may imply that the only crucial concerns in developing an 'optimal orthography' are linguistic ones, it is clear that linguists who have been seriously involved in orthography development believe that socio-cultural concerns are even more important. For example, Pike (1947) writes as a linguist that "a practical orthography should be phonemic. There should be a one-to-one correspondence between each phoneme and the symbolization of that phoneme". But he goes on to state:

> A good alphabet with no [social] motivation will not be read; a poor one with good motivation will allow the absorption of much learning even by people who find reading difficult. (Pike, 1947:215)

Similarly Sjoberg (1966:273) writes the following in conclusion to a discussion of several incidents where linguistic assumptions had to give way to socio-cultural reality in the design of orthographies:

> It is apparent that, in the realm of orthography-making, narrowly linguistic factors are much less important than socio-cultural ones.

In a 1976 panel discussion at the University of Hawaii which dealt with the topic 'What constitutes 'optimal orthography'', Byron W. Bender and Kenneth Rehg, both of whom had been linguistic consultants for the design of orthographies in Micronesia,[2] chose to focus their remarks entirely on social and cultural factors. Among those factors whose importance was discussed by Bender were the following:

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 237-258. *Pacific Linguistics*, C-80, 1984.

(1) Whether the language has ever been written in any form by its speakers, the nature of that form, and the degree to which it is still known and used by all speakers.

(2) Whether the speakers are also speakers of one or more other languages, and whether they read and write those languages.

(3) What uses the speakers will have for reading and writing their language, and what media are apt to be used: print, typing, handwriting, etc.

(4) How strong the desire of the community of speakers is to have their language set to writing, or to have its written form regularised; and what resources they are willing to commit to this purpose: time and energy for education and re-education, and money for publication. If there already exists a body of literature in some other written form, will it be rendered obsolete; will it need to be rewritten; will it create confusion with any revised form to be used in future publication.

Rehg also made the argument that an orthography cannot be accepted unless:

(5) The community acknowledges the authority of the designer(s) and/or implementers of the orthography; and

(6) The community assigns status to knowing the orthography and adjusts to the consequence that not knowing the orthography may result in less status.

Of the many orthographies designed with the assistance of linguists for languages in the Pacific basin, most have been reasonably well accepted. There are some, however, where a retrospective look suggests that some of Bender's and Rehg's points may have been overlooked. In the states of Yap and Truk in Micronesia, for example, sophisticated, linguistically elegant orthographies were proposed and accepted by orthography commissions in the early 1970s. However, the populations of both states were already literate in the (quite unsystematic) orthographies used to translate the Bible. As a result of this (and other factors), the Yapese and Trukese have been quite reluctant to accept the new orthographies,[3] which despite good intentions have been associated by some native speakers with Western imperialism.

This attitude is perhaps especially widespread in the Mortlock Islands, which are part of Truk State and where a language closely related to Trukese is spoken. Despite the fact that several Mortlockese were on the commission that designed the new Truk orthography, many Mortlockese apparently feel it will "steal" from them both their language and culture (Mikeas Olap, p.c.). Under these circumstances it will be very difficult for the new orthography to be given an opportunity to succeed.

In contrast, the new Ponapean and Chamorro orthographies are in large part based on the spelling systems employed in earlier Bible translations. A few phonemic features are now distinguished that were ignored in the earlier systems, no phoneme is now represented by more than one symbol,[4] and the spelling of words has been standardised. Even so, these new systems represent minimal departures from what had existed previously. Both new orthographies are widely accepted. In Ponape, the Bible has now been retranslated into the new system, and in the Northern Marianas, where Chamorro is spoken, there was recently a strong negative reaction to the suggestion of a linguist that the new orthography should be revised so as to indicate all phonemic contrasts.[5]

These examples highlight the importance of Bender's first point. If a speech community already uses a writing system, no matter how linguistically naive that system may appear, it will be very difficult to persuade the community to replace

it with a new one. The importance of Bender's and Rehg's other points will become apparent in Section 3 of this paper. Next, however, I should like to examine some other factors in selecting an orthography.

In an important article, Berry (1958:753), proposed the following three considerations for determining whether an orthography is 'scientifically acceptable':[6]

(1) linguistic: does the alphabet represent the language system economically, consistently, unambiguously?

(2) pedagogical: how does the alphabet achieve the strictly utilitarian aim of economy of time and labor in learning to read and write?

(3) psychological: how far does the alphabet respect the psychological and physiological processes involved in the reading and writing acts?

Like Pike (1947), Berry clearly assumes that a linguistically sound orthography must be founded on the representation of the taxonomic phoneme, on which basis the orthography of, say, modern English would be considered poor.[7] Since the publication of Chomsky and Halle's *Sound Pattern of English* in 1968, however, there has been increasing discussion among linguists and educators of the requirements of a linguistically optimal orthography.

In contrast to Pike, Chomsky and Halle (1968:49) claim that the conventional English orthography is 'a near optimal system for the lexical representation of English words'. The principle for orthographic design that they propose is that 'phonetic variation is not indicated where it is predictable by general rule'. Since the articulation of these two claims, numerous articles and monographs have been written either to demonstrate further the proposed optimality of English orthography (see, e.g., Carol Chomsky 1970, Halle 1969, Weir and Venezky 1968) or to deny that it is optimal (Steinberg 1971, Burling 1976).

Despite the directly opposing claims that Pike and Chomsky and Halle make regarding English spelling, however, the orthographic principles proposed are in many ways more similar than different, and the difference that does exist is more one of linguistic theory than of orthographic theory. Pike, like Chomsky and Halle, assumes that native speakers' tacit knowledge of linguistic rules will enable them to produce the correct phonetic representations when they are presented with the abstract phonemic symbolisations. Of course the phonemic representations posited by Chomsky and Halle are a good deal more abstract than Pike's. It is this difference that causes the differing opinions on English orthography.

Klima (1972) proposes four points of view by which an orthography may be evaluated:

(1) the degree of arbitrariness in the relationship between the orthographic units and the corresponding linguistic units: the less arbitrary the orthography the easier it will be to learn;

(2) the degree of redundancy in the orthographic representation vis-à-vis the linguistic form; the greater the parsimony of the orthography the better...;

(3) the degree of ambiguity in the orthographic representation with respect to the linguistic form represented: the orthography must be suitably expressive;

(4) standardization: one and the same word should not have several spellings, that is, a difference in spelling should represent a difference in linguistic structure.

Like Chomsky and Halle, Klima explicitly assumes that the optimality of an orthography should be determined only with respect to native speakers of the language, and that an orthography that meets native speaker requirements will certainly not be optimal for someone trying to learn the language.

In explaining the 'points of view' stated above, Klima suggests six possible letter-assignment conventions that might be used for English. Four of those conventions are discussed here:

Convention I: Presupposing the dominance of surface phonemes (a term roughly comparable to the Pikean phoneme). In this convention, the writer 'reads up' from the surface phonetic representation of a word toward the systematic representation until he reaches a segment that qualifies as a permissible systematic phoneme. The symbol for that phoneme is then used to represent the segment. In other words, the least abstract segment in a phonological derivation that corresponds to a systematic phoneme is represented by the orthography. Assuming the existence of fairly abstract underlying representations, the resulting orthographic representations of some English words might be chosen as follows (the underlined segments represent the least abstract segments that are permissible systematic phonemes):

| | | | | | |
|---|---|---|---|---|---|
| Underlying Form: | /mis+d/ | /mūv+d/ | /kæ̲nt/ | /rotē̲t+ion/ | /ra̲yd+er/ |
| Derivational Stages: | mist | muwvd | kæ̃n̲t | rotē̲šion | ra:yd̲er |
| | mist | | kæ̃øt | roteš̲øo̲n | ra:yDe̲r |
| Phonetic Form: | [mi̲s̲t] | [mu̲wvd] | [kæ̃t] | [ro̲te̲yš̲n̩] | [ra̲:y̲Dɾ] |
| Orthographic Form: | mist | muwvd | kænt | roteyšon | rayder |

Neither long nor nasalised vowels nor syllabic consonants nor flaps are assumed to be phonemic for English, and as a result they are not represented orthographically. However, /š/, /t/, /d/, and the glides are permissible phonemes, so they are represented.

Convention II: Presupposing the dominance of underlying sound segments. In this convention orthographic representations would be based entirely on the abstract systematic phonemic representations of the segments in the lexical items: the predictable application of a phonological rule would not be reflected in the orthography. Much as the preceding convention closely adheres to that of Pike, this one corresponds to that of Chomsky and Halle. As regards the preceding English examples, they would be written misd, mūvd, kænt, rotētion, and rayder, respectively, with the spelling corresponding directly to the underlying form.

Convention III: Presupposing the dominance of underlying sound segments and morphological boundaries. As there seem to be cases where some phonological rules apply only across boundaries but not within morphemes, under this convention the boundaries would be made explicit. Otherwise, it is identical with Convention II. In English, it would distinguish between, for example, the representations of finger finger and singer sing-er.

Convention IV: Presupposing the transparency of only regular phonological rules. Under this convention, the application of regularly productive phonological rules would not be reflected in the orthography (i.e., the orthographic representation would correspond to the phonological representation before the rules applied). On the other hand, the output of minor or suppletive rules would be reflected. Referring once again to the English examples discussed above, the orthographic representations would be misd, mūvd, kænt, and rayder, as in Convention II, but rotēšon, close to Convention I.

If we assume for the sake of argument the validity of the type of abstract phonological analysis proposed by Chomsky and Halle, it is not at all clear which of Klima's conventions is preferable on solely linguistic grounds. If we turn for a solution to recent studies of the writing and reading process, as Berry suggests, we find a great deal of relevant information, but in the end the situation is not much clearer.

For example, a number of experiments have demonstrated that fluent adult readers make very little use of phoneme-grapheme correspondences in reading. Instead, fluent readers use their semantic and syntactic knowledge to predict the messages they are reading, and where individual words need to be identified, they are recognised as gestalts. Although it is true that fluent readers have the capability of assigning sounds to unfamiliar orthographic forms, the evidence suggests that this is accomplished through syllable recognition and not through an analysis of each individual grapheme. Less conclusive evidence also suggests that writers pay little attention to phoneme-grapheme correspondences as well, but use memorised motor gestalts in the writing of messages (Smith, 1973). Thus, for the fluent reader or writer there is relatively little advantage to an alphabetic orthography over a logographic one which directly represents morphemes and is not phonologically based at all.[8]

For children learning to read and, especially, write, however, the situation may be somewhat different. In a series of studies of children's writing of English before the commencement of formal education, Read (1971, 1973, and 1975) has demonstrated that young children's attention is focussed almost entirely on phonetic characteristics of speech. For example, Read's subjects consistently represented alveolar stops before 'r' as affricates and did not represent nasal consonants at all before voiceless stops, e.g. in camp. (It should also be mentioned that in contrast to this typical behaviour, most of the children selected single representations for the past and present tense suffixes, regardless of the pronunciation.)

If these children's writings reflect the phonological representations in their memories, it seems plausible that they might acquire reading skills more quickly if the conventional orthography also reflected these representations. Furthermore, since it appears that once an orthography has been learned there is no great advantage of one type of orthography over another, such a surface-oriented orthography would presumably not disadvantage an adult reader and writer.

There are, moreover, further considerations to be taken into account. One of the more important is the matter of dialectal variation. An orthography that matches the surface representations of one dialect is certain to conflict with those of another. This problem is somewhat alleviated when a more abstract orthographic representation is selected. A second solution is to let the orthography reflect the variations of individual dialects. This solution makes learning to write relatively easy but in consequence creates additional problems for the reader (no more so, perhaps, than occur in the comprehension of different

dialects when they are spoken — a task that all language users must master). This solution has in fact been effected in Yugoslavia (Lyovin, 1976).

It appears that further theoretical discussion of what constitutes an optimal orthography will only becloud the issue more. In the remainder of this paper I shall turn to an immediate, practical concern: the task begun in 1976 to develop a practical standard orthography for Saipan Carolinian. Although all of the points that I have raised in this introductory section (and other points as well) clearly had relevance to that task, it will be seen that the Carolinians themselves have taken a very eclectic approach to choosing an orthography.

In the next section of this paper, I discuss briefly the linguistic and social information that was most relevant to the selection of a Carolinian orthography. In Section 2, I describe the history of the orthographic decisions made over the last five years, and in the final section I present some final thoughts and conclusions.

## 2. SOCIAL AND LINGUISTIC BACKGROUND

With the exception of the large island of Guam, located approximately 100 miles to the south, Saipan is the largest of the chain of high islands known as the Mariana Islands, which cover the area between 13° and 21° north latitude and 144° and 146° longitude. Although large in relation to its neighbours, Saipan is only sixteen miles long, and four miles wide at its widest point. By far the largest proportion of its population is located in a series of small connected villages along the west coast, which fronts on a large and beautiful lagoon; the east coast consists almost entirely of sheer rocky cliffs. Saipan has served since 1962 as the headquarters of the government of the Trust Territory of the Pacific Islands, which is a United Nations trusteeship consisting now of six governmental districts spread throughout the Marshall and Caroline Islands, or 'political Micronesia'. Saipan also serves as the seat of government of the newly formed United States Commonwealth of the Northern Mariana Islands.

There are approximately 15,000 native inhabitants on Saipan, of whom approximately 11,000 speak the Western Austronesian language Chamorro, which is also spoken throughout the other islands of the Northern Marianas District and on Guam. The remaining 4,000 inhabitants of Saipan speak one of the three Trukic (TK) dialects that together have come to be called Saipan Carolinian, or simply Carolinian (CRL).

Carolinians came to Saipan from islands in the Central Carolines in a series of migrations, possibly commencing as early as the end of the seventeenth century when the indigenous Chamorros were forcibly removed to Guam by the ruling Spaniards, not to return until the middle of the nineteenth century. According to Spoehr (1954), however, the first date that can definitely be set is 1815, when a major typhoon devastated the central Carolines and caused a small group of Carolinians to ask permission of the Spaniards to settle on Saipan. Other migrations followed throughout the nineteenth and early twentieth centuries. According to Trifonovitch (1968), the original home islands of the Carolinians were the atolls of Satawal, Elato, Lamotrek, Puluwat, and Namonuito. Some immigrants have also come from the Hall and Mortlock Islands.

Although the Chamorros have not fared especially well under the dominance of a sequence of metropolitan powers,[9] the minority Carolinians have fared even less well. While Chamorros have had to learn two languages, their own and that of the ruling power, Carolinians have had to learn three. When Chamorros have gone to church they have been able to pray, listen to sermons, take communion, and sing in their native language; Carolinians have had to do these in Chamorro. The Bible was translated many years ago into Chamorro (CHA), but has still not been translated into CRL.[10] Several grammars and dictionaries have been prepared over the years for CHA, beginning with the grammar of Sanvitores in 1668, but to my knowledge CRL has only been described in the short grammar and word list of Fritz (1911). Until 1976 (see below), a Carolinian child entering school was expected to deal immediately with not one but two foreign languages, English and CHA. Teachers would consider him slow, inattentive, and a potential trouble-maker. In these circumstances it is not surprising that Carolinian children have in general performed poorly in school.

Beginning in early 1976, however, and due in large part to the personal efforts of the then director of the Chamorro Bilingual Project, Kit Porter, and of the director of the English programme, Imants Klingbergs, steps began to be taken to provide equal educational opportunities for the Carolinian population. A Carolinian Parent Advisory Council was established, and it immediately requested the Marianas Department of Education to institute a bilingual education programme for CRL, to begin in Fall 1976. Recognising that an important first step in the institution of this programme would be the development of a standard writing system for CRL,[11] The Council chose a young but well respected Carolinian, Jesus Mareham Elameto, to attend the University of Hawaii, where he would receive training at the Bilingual Education Program for Micronesia, but where his first priority would be to work with me on the development of a CRL orthography.[12] The proposed orthography would in turn be presented at the first Carolinian Orthography Convention, which was to be convened on Saipan in July.

## 2.1. Linguistic characteristics of Carolinian

The two major concentrations of Carolinians on Saipan are the villages of Tanapag in the north and Oleai and Chalan Kanoa about eight miles to the south. Tanapag is the home of the people whose ancestors migrated from the atolls of Unoun and Namonuito; however, many of the families here use CHA in their homes, and much of the original Carolinian culture has been modified. In contrast, in the southern villages where the other two dialects are located, CRL is spoken consistently in the home, and there is a strong sense of political and cultural unity.[13]

There is only one significant phonological difference between the two southern dialects: historic *n and *l have merged as /l/ in one (the most populous), while they remain separate in the other. Tanapag speech is quite distinct. In the following chart, Tanapag phonemes are indicated with a T, while the southern village phonemes are indicated with an S.

Table 1: Inventory of single consonant phonemes in Tanapag (T) and southern (S) dialects

| | Labial | | Alveolar | | Palatal | | Retroflex | | Velar | | Glottal | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | T | S | T | S | T | S | T | S | T | S | T |
| Stops [-vd] | p | p | t | t | | | | | | | | |
| [+vd] | $b^w$ | $b^w$ | | | | | | | | g | | |
| Fricatives | f | f | s | | | | ʂ | | x | | | h |
| Nasals | m | m | n | n | | | | | ŋ | ŋ | | |
| | $m^w$ | $m^w$ | | | | | | | | | | |
| Liquids | | | l,r | l,r | | | | ɾ | | | | |
| Glides | w | w | | | y | y | | | (w) | (w) | | |

In addition, both dialects include the following single consonant phonemes in words that have been borrowed from CHA, Jp., or Eng.: b, d, dz, g, s, ts.

All consonants may occur as geminates, at least medially, in all three dialects. For the most part, the geminate variety of each consonant is simply longer and tenser than the single segment. For the following phonemes, however, the phonetic quality of geminate consonants is markedly distinct from that of the single phones:[14]

| Single Phones | Geminate Phones |
|---|---|
| [$b^w$] | [$p^wp^w$] |
| [x,g] | [kk] |
| [ʂ,ɾ] | [ɕɕ] |
| [h] | [ss] |

All consonants are unaspirated, although when geminate the stops often seem aspirated. All obstruents and nasals are obligatorily released word-finally.

Although there is some variation among the different dialects in the vowels used in individual words (indeed, there is significant variation among individual speakers of a single dialect), all three dialects have the same nine vowel phonemes, each of which may occur both short and long:

| | Front | Central | Back |
|---|---|---|---|
| High | i | ɯ | u |
| Mid | e | ə | o |
| Low | æ | a | ɔ |

The history of CRL (and of all other TK languages) has included several important rules affecting vowels. These historical rules have given rise to important morphophonemic relationships in the synchronic grammar of all three CRL dialects, which must be dealt with by an orthographer. These include the so-called process of 'compensatory lengthening' (discussed in Rehg, this volume), and processes of vowel deletion and assimilation.

The effect of the lengthening rule has been to lengthen the vowel of phonetically monosyllabic nouns and pronouns that do not also contain geminate consonants. Thus no noun has less than two mora in its surface free form. Lengthening does not apply to suffixed nouns, but there are some nouns where the vowel is inherently long (most commonly due to the loss of a medial consonant historically), and in these cases the vowel in the suffixed forms remains long.

Forms reflecting the lengthening rule:

| | *house* | *skin* | *sea* | *father* | *road* |
|---|---|---|---|---|---|
| Free noun | iimw | xiil | sææt | saam | aal |
| _____ *of* | imwal | xilil | setil | samal | alal |

Forms with inherent length:

| | *nose* | *song* | *whale* | *man* |
|---|---|---|---|---|
| Free noun | bwoot | kkəəl | rɔɔs | mwææl |
| ____ *of* | ɓwootɯl | kkəəlul | rɔɔsol | mwæælel |

These forms and the ones below also provide synchronic evidence for a historical rule of word-final vowel deletion. The vowels thus deleted are retained before suffixes.

| Gloss | *skin* | *house* | *eye* | *nose* | *head* |
|---|---|---|---|---|---|
| PTK | *kili | *imwa | *mata | *pwauθu | *cimwa |
| Free form | xiil | iimw | maas | bwoot | ṣiimw |
| *my* ____ | xiliy | imwey | mesey | bwootiy | ṣimwey |
| *your* ____ | xilimw | imwɔmw | mesɔmw | bwootumw | ṣimwɔmw |
| *his* ____ | xilil | imwal | mesal | bwootɯl | ṣimwal |

Final vowels also surface before directional and object suffixes and in several other environments: (1) in nominal compounds; (2) in disyllabic reduplication; and, (3) in rapid speech, before demonstrative and adverbial enclitics.

| Loss of vowel | | Vowel retention | |
|---|---|---|---|
| mwææl | *man* | mwæælexaçç | *good man* |
| ææt | *child* | attexit | *boy* |
| xit | *mall* | xitixiit | *little* |
| roŋ | *hear* | roŋoroŋ | *hear* |
| ṣɔɔbwut | *woman* | ṣɔɔbwuto we | *the woman* |
| aramas | *person* | aramasa kkawe | *those people* |
| takk | *finished* | takka mwo | *finished yet* |

Another important rule in the historical development of all three CRL dialects was the loss of single *k before low vowels.[15] Since geminate *kk was not lost, alternations like the following are common:

| | | | |
|---|---|---|---|
| æfiṣi | *like it* | akkafiṣ | *love* |
| abwas | *call* | akkabwas | *shout* |
| apasa | *say it* | kkapas | *word, speech* |
| aŋi | *eat it* | akkaŋi | *eating it* |
| ææmi | 2nd pl ind pron | -xemi | 2nd pl obj pron |

The last morphophonemic alternations in CRL to be discussed in this section are those that occur among the preverbal subject pronouns and aspect markers in the verb phrase. The basic morphemes for these forms are:

| Subject Pronouns | | Aspect Markers | |
|---|---|---|---|
| i | 1st sg | bwe | future, irrealis |
| u | 2nd sg | a | perfective, change of state |
| e | 3rd sg | bwele | immediate future |
| si | 1st pl incl | se | perfective negative |
| ay | 1st pl excl | te | *lest*, imperative negative |
| aw | 2nd pl | ssɔbw | future negative |
| re | 3rd pl | sææl | *unaccomplished* |
| | | saa | *no longer* |

When the subject pronouns are followed by the perfective aspect marker -a, most of the resulting forms are predictable, but the third person forms are not: [aa] 'third person singular perfective'; [raa] 'third person plural perfective'. Here the vowel of the subject pronoun has presumably assimilated to the following vowel. In forms involving the negative markers se and te, following vowels assimilate to preceding ones:

| | |
|---|---|
| i + se → [isi] | si + se → [sisi] |
| u + su → [usu] | ay + se → [aysi] |
| e + se → [ese] | aw + se → [awsu] |
| | re + se → [rese] |
| u + te → [utu] | e + te → [ete] |

All of thee matters created problems in the development of an effective orthography.

## 3. STAGES IN THE SELECTION OF AN ORTHOGRAPHY FOR CAROLINIAN

### 3.1. Naive orthographies

In addition to conducting research into the phonology, morphology, and syntax of Elameto's (southern) dialect of CRL prior to drafting a proposal for a standard orthography, we also systematically examined a representative sampling of nonstandardised writings by adult Carolinians. Although the number of writings collected was not as great as we might have wished, and although the samples differed considerably in some respects — particularly in the representations of the vowels — certain generalisations can be made.

1. Distinctive vowel length was almost never written.

2. Geminate consonants were written occasionally, but normally only in medial position. Occasionally some writers seemed aware that a word contained gemination but were uncertain where to indicate it, e.g. loll for /llɔl/ *in it.*

3. Velarised consonants were indicated by digraphs including either -w or -u, but only prevocalically when the following vowel was nonround. Word-finally or before rounded vowels, velarisation was almost never indicated: lib for /liibʷ/ *hole* but libual *hole of;* muen for /mʷææn/ *man,* but schim for /ʂɨɨmʷ/ *head;* imwei *my house,* but imom /imʷɔmʷ/ *your house.*

4. Only the five vowel letters of the Roman alphabet were used to indicate the nine distinctive vowels. Thus i might correspond to both /i/ and /ɯ/, u to both /u/ and /ɯ/, e to /e/, /æ/, and /ə/, a to /a/, /æ/, and occasionally /ɔ/, and o to /o/ and /ɔ/. Occasionally digraphs were used, such as eu for /ə/ and iu for /ɯ/, but then they were not contrasted with the corresponding vowel-glide sequences.

5. No attention was paid to morphophonemic regularities such as the predictability of the vowel qualities before possessive suffixes; instead the writers' attention was focussed on the surface phones.

6. The subject pronoun was almost invariably attached to the part of the verb phrase that immediately followed, i.e. the negative marker, the aspect marker, an aspectual adverb, or the verb itself.

7. Singular determiners were usually attached to the preceding noun; plural determiners were usually separated from them.

8. The longer object pronouns were occasionally separated from a preceding verb stem, while the shorter pronouns were invariably attached, e.g. e weriei *he sees me* vs. e uri ghish *he sees us* (incl.).

9. Directional suffixes were always attached to the preceding verb.

10. Morphemes phonologically assimilated to other morphemes were occasionally not written at all, e.g. ito for /i+itto/ *I come* (phonetically [iitto]).

11. Although initial and medial /w/ were sometimes written as w, they were more commonly represented as u, and final /w/ was always written as u. The glide /y/ was always represented as i: ua /waa/ *canoe;* motiu /mɔɔttiw/ *sit down;* keue or kewe /kkewe/ *those;* ial /yaal/ *his* (general object); mai /mææy/ *breadfruit.*

It is possible that many of these characteristics can be traced to the influences of other orthographic systems with which Carolinians are familiar, and particularly to the CHA system. For example, CHA does not have phonemic vowel length distinctions; as a result there is no need for an orthographic mechanism to reflect long vowels. Similarly, while CHA has geminate consonants (in the Saipan dialect), they only appear medially, and traditionally were not reflected in the orthography. (The newly standardised orthography, however, does reflect them (Topping, 1973).) CHA also has no counterpart to the CRL velarised consonants, but does have initial Cw- clusters, where the glide has traditionally been reflected by either w or u. Finally, although CHA has six phonemic vowels, only the five Roman letters are used to reflect them, with a corresponding to both /a/ and /æ/. Given these facts, the first four of the above generalisations regarding CRL writing may be explained; perhaps Carolinians tried to apply the orthographic conventions of a familiar but quite different language to the representation of their own language.

The text below is from one of the more systematic southern dialect writings. Phonetic transcriptions and morpheme glosses are also given:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Igha | iya | tiwo | raghi | iya | gukulai | bue | ibue | tata | wal |
| [ixa | iya | tiwoow | ræxiy | iya | xukkuleey | bʷe | ibwe | təətæ | wɔɔl |
| *When* | *I*-perf. | *nine* | *year-my* | *I*-perf | *redup.-learn-it* | *that* | *I*-fut | *climb-up* | *on-it* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| torongkol | meai. | Lughul | imememem | nge | eor | elu | torongkol |
| toroŋkool | mææy | luxul | imwememem | ŋe | eyoor | eluuw | toroŋkool |
| *trunk-of* | *breadfruit* | *outside-of* | *house-our* | topic | *it-exist* | *three* | *trunk-of* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| meai. | Ruwo | meaikal | relo | | arol | immual | rushai |
| mææy | Ruwo | mææy | kkaal | re lo | arol | imwal | ruɯʂay |
| *breadfruit* | *two* | *breadfruit* | pl-*that* | *they stay* | *near-of* | *house-of* | *two*-animate |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| primal | tatai. | Ikkal | meaika | ai | afashi | bue | aibue | |
| priimaal | tæætææy | ikkaal | mææy | kka | ay | æfiʂiy | bwe | ay bwe |
| *cousin-of* | *father-my* | pl-*that* | *breadfruit* | pl-*that* | *we* | *like-it* | *because* | *we*-fut |

tata wal bueigha re arab ngeli im me re lang segi
təətæ wɔɔl bwe ixa re arap ŋæliy iimw me re llaŋ sæŋiy
*climb-up on-it because-then they next to-it house and they tall from-it*

lugal im.
luuxal iimw.]
*centre-of house*

This text exemplifies most of the above generalisations about naive CRL orthographies and also provides the most common written representations of the following CRL consonants:

| Phoneme | Naive written representations |
|---|---|
| /t/ | t |
| /s/ | s |
| /ṣ/ | sh |
| /x̣/ | gh, sometimes g |
| /m/ | m |
| /ŋ/ | ng, sometimes g |
| /l/ | l |
| /r/ | r |
| /kk/ | k, sometimes kk |

In other passages, /p/ was normally represented as p, /f/ as f, /c̣c̣/ as ch, or tch. In Tanapag samples, /n/ was always written as n and /h/ as h, but /r̥/ was most commonly not distinguished in writing from /r/. Occasionally it was written as rh.

## 3.2. Recommendations presented to the first orthography convention

After extensive discussion, it was decided that the most important characteristic for the proposed orthography would be close approximation to the practices observed in the writing samples that had been collected. There were two reasons for this decision. First, since the schools on Saipan would open less than one month after the conclusion of the convention, it was felt that the orthography should be as familiar as possible for the teachers to be able to use it. Second, it was assumed that if the proposal were close to familiar practice, the chances of it being accepted would be enhanced.

Accordingly, the following specific recommendations were made:

1. No distinction should be made between any short and long vowels.

2. No distinction should be made between single and geminate consonants. Optionally, geminate consonants could be distinguished medially.

3. Velarised consonants should be written with digraphs ending in -w because -u can too easily be mistaken for a vowel; otherwise, regular practice in the symbolisation of these consonants should be continued: the -w should not be written word-finally or before rounded vowels.

4. All nine phonemic vowels should be distinguished by individual symbols. To avoid ambiguity with the diphthongs, digraphs should not be used.

5. The orthography should reflect surface phonemes; morphophonemic alternations should be fully expressed despite their predictability.

6. The subject pronoun should be separated from all following morphemes except the negative and aspect markers, to which it should be attached.

7. Singular and plural determiners should be treated in the same way, whether attached to the noun or separated.

8. Object pronouns should all be attached to the preceding verb.

9. Directional suffixes should not be attached to the verb, but the homophonous aspectual particles should.

10. All morphemes should be distinguished in formal writing.

11. The orthography should reflect dialectal variations; there should be no attempt at selecting a standard dialect at this time.

The orthographic symbols that were recommended to the convention were as follows:

| Grapheme | Phoneme | Reason |
|---|---|---|
| a | /a/ | Standard practice. |
| e | /e/ | Standard practice. |
| i | /i/ | Standard practice. |
| o | /o/ | Standard practice. |
| u | /u/ | Standard practice. |
| á | /æ/ | To agree with Trukese (TRK) and Puluwatese (PUL) orthographies and thus assist in communications among the TK communities. |
| é | /ə/ | To agree with TRK and PUL. |
| ó | /ɔ/ | To agree with TRK and PUL. |
| ú | /ɯ/ | To agree with TRK and PUL. |
| f | /f/ | Standard practice. |
| h | /h/ | Standard practice in Tanapag. |
| s | /s/ | Standard practice. |
| sh | /ṣ/ | To assist in the later reading of Eng. Also common practice. |
| gh | /x/ | Standard practice among most Carolinians. |
| k | /kk/ | Most frequent practice. |
| l | /l/ | Standard practice. |
| m | /m/ | Standard practice. |
| mw, m | /mʷ/ | See Point 3 above. |
| n | /n/ | Standard practice. |
| ng | /ŋ/ | Standard practice. |
| p | /p/ | Standard practice. |
| pw, p | /bʷbʷ/ | Standard practice. Geminate /bʷ/ is almost always devoiced. See Point 3 above. |
| bw, b | /bʷ/ | See Point 3 above. |
| r | /r/ | Standard practice. |
| rh | /ṛ/ | For similarity with the symbol for /ṣ/, to which it regularly corresponds. Also common practice in Tanapag. |
| ch | /ç/ | To assist in later reading of English. Also to agree with TRK and PUL. |
| t | /t/ | Standard practice. |
| w | /w/ | Standard practice, except word-finally. Need for consistency. |
| i | /y/ | Standard practice. |

## 3.3. The decisions of the Convention

The Carolinian Orthography Convention met for approximately four hours a day from July 26 to August 4, 1976, at the conference room of the Marianas Department of Education. Twenty-two highly respected members of the Carolinian community attended the Convention, representing the three CRL dialects and the major political and social organisations of the community. In addition to the Carolinian participants, representatives of the Marianas Department of Education and of the Chamorro Orthography Committee also attended on an irregular basis. The Convention was chaired by Jesus Elameto. I served as a non-voting linguistic consultant.

The first task undertaken by the Convention was to determine whether a standard dialect could be decided on for the purpose of standardising the spellings of CRL words in schools and in official documents. Although there was some support for selecting such a dialect, particularly among speakers of the southern dialect containing both /n/ and /l/, eventually it was decided that such a far-reaching decision should not be made at that time. Accordingly, it was decided that the actual task of the Convention would be to select a standard system of representing orthographically the pronunciations of all three dialects, and that Carolinians should use that system to reflect their own individual pronunciations. Two corollaries to this decision were that school teachers were not able to give standard spellings of words to pupils and must instead assist each pupil to learn how to represent his or her own speech, and that textbooks for the early grades of the bilingual programme must be printed in three versions, one for each of the major dialects.

Once that decision had been made, the Convention members turned to a consideration of the alphabetic symbols that would be used in the orthography. For the majority of the symbols that were chosen, discussion was very brief. The symbols decided on with no debate were the vowel letters a, e, i, o, u, and w. The selection of symbols for the remaining four phonemic vowels took longer: while every conference participant wanted these vowels distinguished, some older participants suggested the use of the umlaut diacritic, and other participants wanted to explore the possibility of using digraphs. Eventually, however, the probable advantage of modelling the symbols after those used by Truk State proved persuasive, and the recommended symbols á, é, ó, and ú were chosen.

The next issue to be taken up was the selection of symbols for the velarised consonants /bʷ/ and /mʷ/. Most of the participants were under the misapprehension that the initial consonants in such words as /mʷææl/ *man* and /bʷeel/ *taro patch* were in fact clusters consisting of either /m/ or /b/ and the glide /w/, and that the symbols already decided on, with the addition of b, were already adequate to represent these phonemes. From this point of view, the recommendation presented to the Convention made little sense; the participants seemed unaware consciously of the presence of velarised consonants word-finally or before rounded vowels. After lengthy discussions of possessive paradigms for nouns ending in velarised consonants, the participants came to appreciate their unitary status. Once this was understood, the Convention unanimously decided to represent these consonants with the same symbols, regardless of where they appear in words. The symbols chosen were mw and bw.

One of the two most controversial decisions of the Convention came next. More than five different symbols were suggested for both of the two retroflex obstruents. The participants were adamantly opposed to using the sh and ch as recommended, because, they said, the sounds in CRL were markedly different from the phonemes of Eng. that are represented by those symbols. After discussion, the symbol sch was selected to represent the spirant /ṣ/, and the

symbol tch was eventually selected for the geminate stop, largely because it seemed to match the earlier choice.

The question of whether to represent consonant gemination was then discussed, and it was decided that geminates should in fact be symbolised in initial and medial positions by doubling the consonant letter. Where a single consonant was represented by a digraph, only the initial letter of the digraph was to be doubled, e.g. bbw, mmw, nng, ssch, ggh. Devoiced /bʷbʷ/ was to be represented with pw. For final geminates, the Convention first decided to use the doubled symbols and then, on the last day, this decision was rescinded and it was decided to use single consonant letters, including single k for /kk/. Once the decision regarding symbolisation of the geminates had been made, it took very little further discussion to decide that long vowels should not be distinguished.

Although the letter w had already been chosen to represent /w/, at the end of the Convention it was decided that w should be written only in initial position or between unrounded vowels. In final position /w/ was to be written with u. It was further decided that /y/ should be represented by i in all positions.

The Convention also decided that the following letters could be used in the spelling of foreign words and proper nouns: b, c, d, g, j, k, n, q, v, x, y, z.

Decisions made regarding when to write two morphemes together as one word and when to write them separately as two words generally emphasised the need for young school children to be able to recognise words quickly and easily. It was expected that in informal writing such as letters Carolinians would continue to write several morphemes together as single words. The decisions about word division can be summarised as follows:

1. Possessive suffixes should always be written attached to the preceding nouns.

2. All determiners should be written separately from the nouns they modify.

3. Subject pronouns should be written attached to following aspect markers, but should be written separately from any other element in the verb phrase.

4. All object pronouns should be attached to the preceding verb stem.

5. Both aspectual particles and directional suffixes should be written separately from the preceding verb.

With very little debate it was decided on the last day of the Convention that foreign words which have become assimilated into CRL should be spelled the way they are pronounced by Carolinians (e.g. resituran < Eng. *restaurant*, bwola < Sp. *bola*), but that unassimilated borrowings should be spelled the same as in their original language and should also be underlined to indicate their foreign origin (education, president, commonwealth), that the new orthography should not influence the spelling of CRL names, such as Jesus, and that the system of punctuation to be used should be identical with that of Eng. so as not to confuse schoolchildren, except that the first person pronouns ngang and i should only be capitalised at the beginning of a sentence.

The following table summarises the Convention's decisions regarding the recommendations made to it.

| Recommendations | Convention Decisions |
|---|---|
| 1. Vowel length distinctions should not be represented. | 1. Agreed. |

| | |
|---|---|
| 2. Geminate consonants should not be distinguished from single consonants. | 2. Initial and medial geminates should be distinguished. |
| 3. Velarised consonants should be written Cw; they should not be distinguished before round vowels or at the ends of words. | 3. Agreement on the first recommendation; the consonants should be distinguished wherever they occur. |
| 4. All nine phonemic vowels should be separately represented. | 4. Agreed. |
| 5. The orthography should reflect morphophonemic alternations. | 5. Agreed. |
| 6. The subject pronoun should be separated from all morphemes except negative and aspect markers. | 6. Agreed. |
| 7. Singular and plural determiners should be written in the same way, whether attached or not attached. | 7. Both singular and plural determiners should be separated from the noun. |
| 8. Object pronouns should all be attached to the verb. | 8. Agreed. |
| 9. Aspectual particles should be attached to the verb; directional suffixes should be separated. | 9. Both aspectual particles and the directional suffixes should be written separately. |
| 10. All morphemes should be distinguished in formal writing. | 10. Agreed, |
| 11. The orthography should reflect dialectal variations. | 11. Agreed. |

It was in conclusion decided to convene a Second Carolinian Orthography Convention two years later to review the first Convention's decisions.

## 3.3. Decisions of the Second Carolinian Orthography Convention

The second Convention was held in January, 1978, and was attended by almost all of the previous participants plus about 20 new ones, including two Carolinian legislators. In general, this meeting approved the earlier decisions, but it also made five new and in some ways surprising recommendations. First, in large part due to the suggestions of teachers and curriculum writers who had been using the new orthography, it was recommended that vowel length should be distinguished and that directional suffixes should be attached to the preceding verbs. Representation of vowel length was urged so as to minimise homographs, and also because, as one participant said, CRL has long vowels, so they should be written.

The third decision was not to write vowels that surface before enclitics, but to continue to write them before suffixes and in compounds. The fourth was to represent the phoneme /y/ as y in all positions, and having made that decision, the Convention also chose to represent final /w/ as w, rather than as u.

The decision to write y was arrived at largely by accident. The Convention was exploring the spellings of reduplicated words and happened to examine a relatively rare pattern of reduplication in verbs commencing with /i/, such as /iṣ/ *write*. Although most speakers either reduplicate that verb ikkisch or do not reduplicate it at all, others insert a prothetic [y] and treat the glide as an initial consonant, in which case the reduplicated form is [yiyyiṣ].[16] Under the former orthographic conventions, that word would be written iiiiisch, a spelling that was met with horror by the Convention. As a result, the letter y was introduced.[17]

## 4. CONCLUSIONS

More than three years have passed since the last Orthography Convention was held, and it appears that its recommendations will stand. The Bible is now being translated into CRL using the new orthography, and a short book of hymns and prayers has been rewritten in it. Perhaps equally important, a Carolinian-English dictionary was to appear in late 1982 which will further codify the new system.

Most CRL writings that I have had a chance to observe over the last few years, by both adults and children, adhere fairly closely to the decisions that were made, but there are still a few common divergences. Most common, perhaps, is the failure to represent predictable vowel length (e.g. as the result of the process referred to as 'compensatory lengthening'). In addition the subject pronoun is often written attached to a following adverb or verb stem, especially if that word is a short monosyllable (e.g. ghal *habitually*, kke *repeatedly*, ló *go*, lo *stay*, má *die*). I suspect that these 'errors' will continue to be common, and that they will probably become the preferred spelling pattern eventually.

If so, it will be in keeping with the process by which this orthography has been selected (i.e. not 'developed') by the Carolinian people. From a linguistic point of view, the process of selecting the CRL orthography seems at best haphazard, and in many ways the result itself fails to conform with what theory might prefer. Indeed, it fails to meet many of Berry's and Klima's requirements for an optimal orthography. It is not standard, in that each word is to be spelled according to the writer's pronunciation. It is not parsimonious, in that many predictable alternations are represented phonetically. In practice it is not even consistent. But I do not believe that this matters.

Although linguistic input has been provided by myself and by others upon request, both the impetus for the orthography and the decisions made have come directly from the Carolinians themselves. The orthography is now both written and read by a large majority of Carolinians — from the age of five to 65 — and several elderly Carolinians have expressed to me their pride in their grandchildren's fluency in the written language. Further modifications of the present system may yet occur, but I think there is little question that the Carolinian orthography will remain vital and important in the community.

## NOTES

1. This paper is an extensive revision and updating of a paper co-authored with Jesus Mareham Elameto which appeared in 1977 in the University of Hawaii *Working Papers in Linguistics*. It has benefitted greatly from comments and suggestions from Rosa Roppul Warakai, Hiroshi Sugita, Byron W. Bender, Jean Kirschenmann, Bonnie Davis, and especially Ken Rehg, to all of whom I am indebted. I retain full responsibility, however, for any errors of commission or omission.

2. Bender assisted in the design of the Marshallese orthography, while Rehg was a consultant in the design of the standard Ponapean orthography.

3. This attitude has also been a factor in the relative lack of acceptance of the vernacular language education programmes in those states.

4. The earlier Chamorro orthography, which was based on Spanish, had represented the phoneme /k/ by the symbols k, c, and qu. These last two symbols have now been dropped from the orthography. My statement needs to be modified slightly in the case of Ponapean, where /w/ is still represented by w initially and medially but by u finally.

5. The new Chamorro orthography fails to distinguish the phonemes /a/ and /æ/. Following Spanish tradition it also uses the symbol y to represent /dz/, despite recent English loans with the phoneme /y/. The linguistic consultant argued that all four phonemes should be distinguished.

6. Although his paper focussed on the three concerns listed, Berry was also aware of the importance of social or socio-cultural factors in developing an acceptable orthography:

> ... an alphabet is successful in so far and only in so far as it is scientifically and socially acceptable. The two interests often conflict and it would be a fallacy to assume, as it sometimes is done, that the choice of an orthography can be determined solely on grounds that are linguistically or pedagogically desirable....there are factors in the social situation...which merit the linguist's attention and for which he should be prepared to condition his science to circumstances and seek a compromise. (Berry, 1958:752-753)

7. In support of his belief, Pike stated, "Children seem to require two or three times as long to learn to read English as comparable children do to learn to read Spanish, which is written unambiguously." (Pike, 1947:208)

8. In fact, Davis (1977) argues persuasively that the only ultimate advantage for adults that an alphabetic orthography may have over a logographic one like that for Chinese is that the smaller number of symbols in an alphabetic orthography can more easily be printed by the instruments of modern technology, such as the typewriter or computer. As technology advances, however, this advantage may vanish. Steinberg (1977) uses evidence from his study of Japanese children to draw the conclusion that a logographic orthography is in fact easier to learn than a phonologically based one.

9. Spain ruled the Northern Marianas until 1899, followed by Germany (1899-1914), Japan (1914-1944), and the United States.

10. A small CRL hymn book was printed in the early 1970s. The Bible is also now being translated into CRL.

11. An attempt had been made in the early 1970s to develop an orthography, dictionary, and grammar for CRL. However, despite the efforts of the Carolinian researcher, Lino Olopai, that project did not come to fruition.

12. Prof. Roderick Jacobs of the University of Hawaii had offered a Linguistics Field Methods course in the Fall of 1975 which had dealt with one dialect of CRL. The informant for that class, Teresa Taitano, returned to Saipan in January, 1976, to serve as project director for the new Carolinian Bilingual Project. Jacobs continued to coordinate research into the CRL language at the University during Spring, 1976, and it was through him that I became involved in work on CRL. It is likely that without Jacobs' initial efforts on behalf of the Carolinians much of the recent progress would have occurred at a much slower pace, if at all.

13. During the last two years political consciousness has been increasing in Tanapag as well. Especially among the older people of the village a concern has developed that the next generation will completely assimilate to CHA culture unless something is done quickly.

14. In these cases, the geminate phonemes more closely retain the quality of the reconstructed historical segments. For example, both Tanapag /g/ and Southern /x/ are derived historically from PTK *k. Similarly, Tanapag /h/ derives from PTK *t (> pre-CRL *s), and Tanapag /ṛ/ and Southern /ṣ/ almost certainly derive from a PTK retroflex stop. (See Jackson, this volume.)

15. The rule is presumably not productive, as recent loans from CHA and Jp., for example, do not undergo it.

16. The -kk- of ikkisch is thought to have developed by analogy with the reduplication pattern of forms in initial *ka, the *k of which was later lost (see the preceding disucssion of the loss of single — but not geminate — *k before low vowels). A parallel development for TRK is described by Goodenough (1963) under the heading 'The -kk- infix'.

17. Although the Second Convention agreed to write y in all places where the phoneme /y/ occurs, the Department of Education has since sent out a directive that it be written i when it occurs immediately before a consonant. The main reasons for this appear to be aesthetic.

## BIBLIOGRAPHY

BENDER, Byron W.

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

BERRY, Jack

1958 The making of alphabets. In *Proceedings of the VIII International Congress of Linguists, Oslo,* 752-764(770). Oslo: University Press.

BURLING, Robbins

1976 Poor spellers of the world, unite! *Learning (March)*:76-78.

CHOMSKY, Carol

1970 Reading, writing, and phonology. *Harvard Educational Review* 40/2: 287-309.

CHOMSKY, Noam and Morris HALLE

1968 *The sound pattern of English.* New York: Harper & Row.

DAVIS, Bonnie (Bonnie Davis ARAMAN)

1977 Who needs an alphabet? *Working Papers in Linguistics, University of Hawaii* 9/1:11-18.

FRITZ, Georg

1911 *Die Zentralkarolinische Sprache: Grammatik, Übungen und Wörterbuch der Mundart der westlich von Truk liegenden Atolle, insbesondere der Saipan-Karoliner.* Berlin: G. Reimer.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The proceedings of the Ninth Pacific Science Congress, 1957,* vol.3:77-86.

GOODENOUGH, Ward H. and Hiroshi SUGITA

1972 Final report of the conference to decide on an orthography for the Truk language. Mimeo. University of Hawaii.

HALLE, Morris

1969 Some thoughts on spelling. In Kenneth S. Goodman and James T. Fleming, eds *Psycholinguistics and the teaching of reading,* 17-24. Newark: International Reading Association.

JACKSON, Frederick H.

1976 Summary of decisions reached by Carolinian orthography committee. Mimeo.

1983 Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. See pp. 259-280 in this volume.

KLIMA, Edward S.

1972 How alphabets might reflect language. In James F. Kavanagh and Ignatius G. Mattingly, eds *Language by ear and by eye: the relationships between speech and reading,* 57-80. Cambridge, Mass.:MIT Press.

LYOVIN, Anatole

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

PIKE, Kenneth L.

1966 *Phonemics.* (10th printing.) Ann Arbor: University of Michigan Press.

READ, Charles

1971 Pre-school children's knowledge of English phonology. *Harvard Educational Review* 41/1:1-34.

1973 Children's judgments of phonetic similarities in relation to English spelling. *Language Learning* 23/1:17-38.

1975 *Children's categorization of speech sounds in English.* Urbana: National Council of Teachers of English.

REHG, Kenneth L.

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

1983 The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See pp. 52-58 in this volume.

SJOBERG, Andrée G.

1973 Socio-cultural and linguistic factors in the development of writing systems for preliterate peoples. In William Bright, ed. *Sociolinguistics,* 260-276. The Hague: Mouton.

SMITH, Frank

1973 *Psycholinguistics and reading.* New York: Holt, Rinehart and Winston.

SPOEHR, Alexander

1954 *Saipan: the ethnology of a war-devastated island.* Chicago: Natural History Museum.

STEINBERG, Danny

1971 Would an orthography based on Chomsky and Halle's underlying phonological representations be optimal? *Working Papers in Linguistics, University of Hawaii* 3/3:2-18.

1977 Learning *kanji* is easier than learning *kama*. *Working Papers in Linguistics, University of Hawaii* 9/1:19-28.

TOPPING, Donald M.

1973 *Chamorro reference grammar.* Honolulu: University Press of Hawaii.

TRIFONOVITCH, Gregory J.

1968 Notes on Saipanese Carolinian grammar. MS

WEIR, Ruth H. and Richard VENEZKY

1968 English orthography — more reason than rhyme. In Kenneth S. Goodman, ed. *The psycholinguistic nature of the reading process,* 185-199. Detroit: Wayne State University Press.

# REFLEXES OF PROTO-OCEANIC IN THE TRUKIC LANGUAGES OF MICRONESIA

Frederick H. Jackson

## 1. INTRODUCTION[1]

The Trukic (TK) languages are spoken by approximately 45,000 people who make up the populations of the large majority of the inhabited high islands and atolls in the Caroline Islands of Micronesia,[2] and by the 4,000 descendants of recent migrants from the central Carolines who now reside on the island of Saipan in the Northern Mariana Islands.[3] Excluding Saipan, the area covered by islands where TK languages are spoken indigenously extends almost 1,500 miles from 132° - 154° E. longitude, and from 3° - 10° N. latitude. Quackenbush (1968:2) has estimated that the number of islands where the TK languages are spoken is at least sixty-six. The number of distinct and separate 'languages' has been estimated at as few as three (Bender 1971) to as many as thirteen (Quackenbush 1968). For the purposes of this paper, we shall assume at least seven distinct languages, which are listed here together with the sources from which data have been obtained.[4]

| | |
|---|---|
| Lagoon Trukese (TRK) | Goodenough and Sugita 1980;<br>Sugita forthcoming;<br>Dyen 1965 |
| Ulithian (ULI) | Sohn and Bender 1971;<br>Elbert 1947;<br>personal notes |
| Pulo Ana (PUA) | Oda 1977;<br>Capell 1969 |
| Mortlockese (MRT) | personal notes |
| Puluwatese (PUL) | Elbert 1972;<br>Elbert 1974 |
| Satawalese (STW) | personal notes |
| Woleaian (WOL) | Sohn 1975;<br>Sohn and Tawerilmang 1976;<br>personal notes |

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 259-280. *Pacific Linguistics*, C-80, 1984.

Although two separate and distinct languages may prove to be subsumed here under the label MRT, and the language of Tobi (TBI) may also prove to be distinct from PUA, all evidence indicates that the above seven languages are representative of TK, and that accurate reconstructions may be made from their testimony.

It is clear phonologically, morphosyntactically, and lexically that the TK languages form a distinct subgroup of Oceanic. The languages with which they are most closely related are almost certainly Kiribati (Gilbertese) (KIR), Ponapeic (PP), Marshallese (MRS), and probably Kosraean (Kusaiean) (KSR), although the status of that language is somewhat problematical.[5] These languages, together with TK, are referred to as nuclear Micronesian (MC) (Bender 1971).

It is not at all clear, however, which languages or language groups outside of Micronesia are most closely related to TK. Pawley (1977) has tentatively assigned MC (and, thus, TK) to Remote Oceanic, almost solely on the basis of somewhat ambiguous data from KIR. Other evidence, however, suggests other groupings. (See Jackson, in preparation, for discussion.)

In the following section of this paper I briefly present the phonological system reconstructed for Proto-Trukic (PTK); in the next section this system is related to Proto-Oceanic (POC), with a discussion of problematic reflexes. The final section offers some tentative conclusions.

## 2. PROTO-TRUKIC PHONOLOGY

Table 1 presents the consonant correspondences among eight TK languages. Phonemes are represented using the orthography of the standard reference works for each language, with phonetic transcriptions added as necessary in square brackets.

As the table demonstrates, most of the TK reconstructions are quite straightforward, the only problematic cases being those symbolised as *k, *c, and particularly *t and *θ. PTK *k (rather than, e.g., *x) is reconstructed because all those languages which show single [x] also show geminate [kk], and because all MC cognates also show [k]. A post-alveolar stop *c is reconstructed because all reflexes except PUA s are either retroflex or palatal, and because those languages which do not exhibit a stop in nongeminate reflexes do so when the reflex is geminate, as with *k.

The phoneme *t is reconstructed largely on the basis of external evidence (as will be seen, PTK *t regularly reflects POC *t), but also in view of the t reflex before a in WOL, ULI, and PUA. PTK *t having been reconstructed for this set of correspondences, a different reconstruction must be made for the set ULI d[θ], PUA-TRK-MRT-PUL-STW-CRL-WOL t which reflects POC *s, *ns. Marck (n.d.) has suggested the reconstruction of a palatal *t$^{y}$ for what I have written *t, and a plain *t for what I have represented as *θ. Although Marck's analysis is partly supported by the fact that t is also the reflex of POC *s in all other MC languages except KIR (where it is r), and although it is probably true that PTK *t (< POC *t) was in fact palatalised,[6] Marck's suggestion does not provide an explanation of ULI d[θ], the development of which would in his analysis require a spirant (POC *s, *ns) to become a stop in PTK and then again a spirant in ULI.

Sohn et al. (1977) have suggested that ULI developed d from a PTK *s under the influence of neighbouring Yapese (YAP), which frequently reflects POC *ns and *s as [θ] (e.g. ULI duud, YAP thuuth < POC *(n)su(n)su *breast*). If this

Table 1: TK consonant correspondences

| PTK | *p | *pw | *f | *t | *c | *θ | *k | *m | *mw | *n | *ŋ | *l | *r | *w |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TRK | p | pw | f | Ø,ʔs | ch[t̪ʃ] | t | k,Ø,[6]s[7] | m | mw | n | ng,n[10] | n | r | w |
| ULI | p | b[βʷ] | f | s,ʔt | c[t̪ʃ] | d[θ] | g[x][9] | m | mw | l | ng | l | r | w |
| PUA | p | pw[βʷ][1] | f | d[θ],ʔt | s | t | k[x][9] | m | mw | n | ng | n | l | w |
| MRT | p | pw | f | Ø,ʔs | sh[ʃ̪][5] | t | k,Ø,[6]s[7] | m | mw | n | ng | l | r | w |
| PUL | p | pw | f | Ø,ʔh,ss[4] | r[ɹ̪][5] | t | k,Ø[6] | m | mw | n | ng | l | ŕ | w |
| STW | p | pw | f | Ø,ʔs | rh[ɹ̪][5] | t | k,Ø[6] | m | mw | n,l[11] | ng | l,n[11] | r | w |
| CRL | p | bw | f | Ø,ʔs | sch[ʃ̪][5] | t | gh[x],Ø[6,9] | m | mw | l | ng | l | r | w |
| WOL | p | b[βʷ][1] | f | Ø,[3]s,[13]t | sh[ʃ̪][5] | t | g[x],Ø[8,9] | m | mw | l[12] | ng | l[12] | r | w |

[1]voiceless stop [pʷpʷ] when geminate

[2]usually before non-low vowels

[3]irregularly before non-low vowels

[4]h → s / when geminate

[5]voiceless stop [t̪ʃt̪ʃ] when geminate

[6]*k > Ø / before low vowels and sporadically before mid vowels

[7]*k > s / ___ i sporadically (less frequent in MRT)

[8]*k > Ø sporadically before low vowels

[9]voiceless stop [kk] when geminate

[10]*ŋ > n / ___ i in most cases

[11]the distinction between *l and *n is apparently collapsing in STW

[12]both *n and *l are [l] singly, [nn] when geminate in WOL

[13]before non-low vowels and irregularly before a

is true, however, the influence must have been something other than a straightforward instance of borrowing. ULI includes many lexical items containing d, possible sources for which are not found in Yapese, and vice versa. Moreover, ULI always reflects POC *s, *ns as [θ], while YAP does not: e.g. ULI taed, YAP daay *sea water* (< POC *tansi(k)); ULI fidig, YAP qu-fi-n *meat, flesh* (< POC *pinsiko). Finally, ULI forms containing d are otherwise normal developments of PTK, showing no influence from YAP.

Several other OC languages reflect *s or *ns as a dental fricative in addition to ULI and YAP, including Fijian (FIJ) and languages of the South-East Solomons. Levy (1980) has in fact reconstructed *ð as the Proto-Eastern Oceanic (PEO) reflex of POC *ns. In these circumstancces, and because TK merges POC *s and *ns (see Section 3 below), it is not implausible that PTK and YAP should separately reflect these consonants as *θ.

All TK languages now have at least seven phonemic vowels (most have nine); vowel length is also distinctive. It is not yet clear whether all seven of the vowels need to be reconstructed for PTK, but it is certain that at least six do (both long and short):

| | | |
|---|---|---|
| *i | *ú | *u |
| *e | | *o |
| | *a | |

The high central unrounded vowel *ú was in pre-TRK an allophonic variant of *u in the environment of a nonround vowel, and of *i before *u. Marck (1977) has also suggested that there was an allophonic rule in PMC and PTK by which *u became *ú after what he terms "front consonants", specifically *p, *t, *m, *n, and *l. However, it is clear that *ú had become distinctive in PTK in, e.g. *núú *coconut* (< POC *niu(R)), and *θúú *bathe* (< POC *(n)su(ŋ)ki).

Other assimilatory vowel allophony must have been present in PTK as well, especially affecting the low vowel *a, which apparently fronted to á [æ] before (C)$\{^{i}_{e}\}$, raised to é [ə] before (C)u, and backed to ó [ɔ] before (C)o.

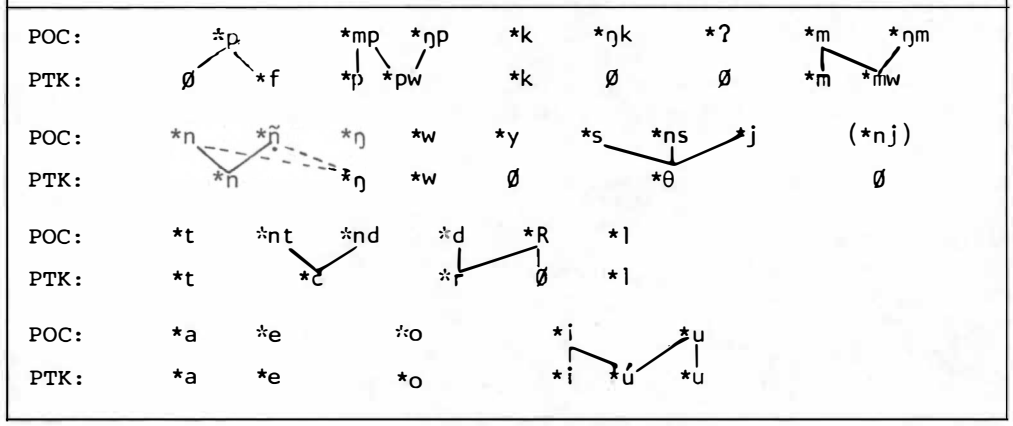
All TK languages also show evidence of a rule of phrase-final vowel devoicing and of the rule that Dyen (1949) has called "compensatory lengthening", by which the first vowel in an unaffixed bimoric noun is lengthened. These rules, too, must be reconstructed for PTK. (See Jackson, 1978, and Rehg, this volume, respectively, for discussion of the two rules.)

## 3. PROTO-TRUKIC REFLEXES OF PROTO-OCEANIC

The POC reconstructions that are compared in this section to PTK have been taken from many sources, including Grace (1969), Blust (1972), Blust (1978), Pawley (1979, n.d.), Lincoln (ms), and Ross (1977). A few new reconstructions are also proposed and supported. In addition, PEO reconstructions from Biggs (1965), Pawley (1972), and Geraghty (1979) have also been examined. All PTK reconstructions are reflected in at least three TK languages, and most are reflected in six or more.

The regular PTK reflexes of POC are shown in Table 2. Details of this pattern, and exceptions to it are discussed below. While all witnesses of such relatively rare POC reconstructions as *ŋp, *ŋk, *nd, *y, etc. are presented, only representative samples of the more frequently witnessed reconstructions are shown. All known exceptions, however, are presented.

Table 2: PTK reflexes of POC phonemes




## 3.1. The labial series of consonants

POC *p is normally lost before round vowels in PTK and reflected as *f elsewhere.

POC *p > PTK ø

| | | |
|---|---|---|
| *pupu *fish-trap* | > | *uu |
| *(n)topu *sugarcane* | > | *tou |
| *pua *fruit,* spherical object | > | *-ua general counting classifier |
| *-ŋapulu *ten* | > | *-ŋaulu |
| *tapu(n)i *afterbirth* | > | *taúú |
| *puti(k) *pull, extract* | > | *útú |
| *punti *banana* | > | *úcú |
| *tapud,Ri *conch* | > | *tawii |
| *nopu *reef fish with poison spines* | > | *neu *stonefish* |
| *(n)sipo *downwards, east* | > | *θiwo[7] |
| *mapo *heal* | > | *mao |
| *yapo *fish-line* | > | *ao |
| *napo *wave* | > | *nao |
| *ñ,nopo *stay, dwell, sit* | > | *noo *stay, dwell* |
| *poñu *turtle* | > | *(w)onú |
| *potu *outside, outward*[8] | > | * otu |

POC *p > PTK *f

| | | |
|---|---|---|
| *faRi- reciprocal prefix | > | *fa- |
| *pale *house* | > | *fale |
| *panua *land, earth, village* | > | *fanúa *island, land* |
| *papine *woman* | > | *faifine[9] |
| *paŋo *snort, blow nose* | > | *foŋo |
| *ŋapa *fathom* | > | *ŋafa |
| *palisi *grass* | > | *faθili[10] |
| *palusa *paddle* | > | *faθula |
| *patu *rock, stone* | > | *fatú |
| *panda(n) *pandanus* | > | *faca |
| *ʔa paRa *shoulder* | > | *afara |
| *pi(n)sa *how much* | > | *fiθa- |

| | | |
|---|---|---|
| *api *fire* | > | *afi |
| *pili(?) *select, choose* | > | *fili |
| *pinsiko *meat, flesh* | > | *fiθúko |
| *pitu?u *star* | > | *fútúú |

Although there are few clear examples, POC *mp is apparently reflected as PTK *p before non-round vowels, and merges with POC *ŋp as the labiovelar *pw before round vowels.

POC *mp > PTK *p

| | | |
|---|---|---|
| *mpaya *bait* | > | *paa |
| *tampi(d)a *bowl* | > | *tapia |
| *mpampa *board, plank* | > | *papa |
| *?e,ampa *mat* | > | *e,apa *baby's mat or cloth* |
| *la(m)pas *big, great* | > | *lapa |
| *mpa(ŋ)kiwak *shark* | > | *pakewa |

POC *mp/*ŋp > PTK *pw

| | | |
|---|---|---|
| *mpua *betel nut* | > | *pwua |
| *(m)ponot *shut, stop up* | > | *pwono |
| *mpo *smell* | > | *pwoe |
| *tampu *ban, taboo* | > | *tapwu |
| *(m)puli *cowry* | > | *pwili |
| *(m)pulu *resin, sap* | > | *pwuli |
| *(m)pono *truth, correctness* | > | *pwono *agree to* |
| *(m)puto *navel* | > | *pwutu[11] |
| *ŋpoŋi *night* | > | *pwoŋi |

There are several exceptions to the above generalisations, however, where PTK appears to reflect nasal grade *mp where only oral grade *p has been reconstructed for POC:

| | | | |
|---|---|---|---|
| POC | *pupu *leak, drip, spill out* | > PTK | *pwuu *flow of liquid*[12] |
| | *pupu *kind of fish* | > | *pwupwu *triggerfish* |
| | *puku *knot, lump, swelling* | > | *pwukua *knee* (?) / *pwuka *knot, navel* (?) |
| | *(n)tupa *derris fish poison* | > | *tupa |
| | *pili *plait, wrap around* | > | *pili *plait, braid* |
| | *peka *faeces, defecate* | > | *pake,a |
| | *patu *knot, excrescence* | > | *pwatu *scar* (?) |
| | *paRata *north-west monsoon* | > | *parata *windstorm* |
| | *pela *dirty* | > | *pwpwelu (??) |
| | *kapu *buttocks* | > | *kapi[13] |

It should also be noted that Geraghty (1979) has suggested an alternative analysis for similar aberrant reflexes of POC *p in FIJ and the South-east Solomons. He has proposed for PEO (and, with qualifications, for POC) a contrast between *p and *v, where previously reconstructed PEO *p corresponds in most cases to his *v. Only three of Geraghty's PEO reconstructions with *p are apparently reflected in TK, but each form reflects a stop, as predicted under Geraghty's hypothesis, rather than a spirant: *papa *board, plank*, *puku *knot*, and *pela *mud, swamp* ( PTK *pwela).

POC *m is reflected as PTK *mw before round vowels and as *m elsewhere, with three apparent exceptions.

POC *m > PTK *m

| | | | |
|---|---|---|---|
| POC | *n,lima *bail* | > PTK | *núma |
| | *manaŋ *supernatural power, wind* | > | *mana |
| | *mea *thing* | > | *mee(-θaa) *what?* |
| | *inu(m) *drink* | > | *únúm- *drink* (vt) |
| | *manu(k) *bird, animal, fish* | > | *manú *creature* |
| | *mata *eye* | > | *mata |
| | *mataku(t) *afraid* | > | *mataku |
| | *matudu(R) *sleep* | > | *matúru |
| | *mate *die* | > | *mate |
| | *masi *breadfruit* | > | *mai |
| | *nsama *outrigger* | > | *θama |
| | *kami 1st pl excl foc pron | > | *kamami/*kami[14] |
| | *kamuyu [EO *kamiu] 2nd pl foc pron | > | *kamii |
| | *dama *torch, light* | > | *marama *moon* |
| | *lumi *fold, crease* | > | *lúmi |
| | *maRa *ashamed* | > | *maa |
| | *lima *five* | > | *lima |

POC *m > PTK *mw

| | | |
|---|---|---|
| *mu 2nd sg poss pron | > | *mwu |
| *mo,utu *cut off* | > | *mwetú *cut off, adopt* |
| *komu *mouthful, gargle* | > | *kumwu *rinse mouth* |
| *n,ñamo *lagoon* | > | *namwo |
| *muta(ʔ) *vomit* | > | *(mw)mwuta |
| *ñamu *mosquito* | > | *namwu |
| *monse *sleep* | > | *(mw)mwee |
| *muʔa *front, precede* | > | *mwoa- |
| *ʔumu *oven* | > | *umwu |
| *mudi *behind* | > | *mwuri |
| *limut *seaweed, moss* | > | *lumwu |
| *ta-molu *man*[15] | > | *tamwoolu *chief* |

The apparent exceptions are two cases where POC *m is reflected as TK *mw before a non-round vowel, and one instance where it is deleted before a round vowel. If in fact the TK forms are cognate, I have no explanation for them.

| | | |
|---|---|---|
| *kima *giant clam* | > | *kamwee[16] (?) |
| *mali(ŋ) *bitter* | > | *(mw)mwale *sour* (?) |
| *(ñ)amu *taste, flavour* | > | *(n)nau *delicious* (?) |

In the five instances where POC *ŋm seems to be reflected in TK, it is realised as *mw four times, and as *m once:

| | | |
|---|---|---|
| *Ruŋma *house* | > | *imwa |
| *ta-ŋmaʔane *man* | > | *mwaane |
| *ŋmata *worm, snake* | > | *mwata *earthworm* |
| *ndaŋma *forehead* | > | *camwa |
| *ŋmalala *cleared ground* | > | *malaala (?) |

## 3.2. The velar series of consonants

POC *k is consistently retained as *k in PTK, while *ŋK and *ʔ are always lost.

POC *k > PTK *k

| | | |
|---|---|---|
| *kapu *buttocks* | > | *kapi |
| *katae *lee side* | > | *katae |
| *toka *land, arrive* | > | *toka |
| *mataku *afraid* | > | *mataku |
| *kainaŋa *clan, descent group* | > | *kainaŋa *clan* |
| *nsa(ŋ)kaRu *reef* | > | *θakau *reef island* |
| *nsake *up, rise* | > | *θake |
| *tuki *strike, beat* | > | *túki |
| *ki(n)ta 1st pl incl pron | > | *kica |
| *masaki *pain* | > | *maθaki |
| *-akini remote trans suffix | > | *-akini |
| *kiñit *pinch, pluck* | > | *kini |
| *koso *husk coconuts* | > | *koθo |
| *(n)soko *caught, captured* | > | *(y)oko |
| *ko 2nd sg pron | > | *ko 2nd sg subj pron |
| *komu *mouthful* | > | *kumwu *rinse mouth* |
| *kutu *louse* | > | *kutú |
| *kuli *skin, bark* | > | *kili |

POC *ŋk > PTK Ø

| | | |
|---|---|---|
| *waŋka *canoe* | > | *waa |
| *-ŋku 1st sg poss pron | > | *-i |
| *(n)suŋki *bathe* | > | *θúú |
| *so(ŋ)ka(r) *cross seat in canoe* | > | *toa (?) |

POC *ʔ > PTK Ø

| | | |
|---|---|---|
| *ʔulapi *parrot fish* | > | *ulafi |
| *paʔu *tie, bind* | > | *faú(faú) |
| *ʔasu *smoke* | > | *aθú |
| *taʔe *not* | > | *tae,i |
| *ʔate *liver* | > | *ate |
| *daʔa *branch* | > | *raa |
| *muʔa *front* | > | *mwoa |
| *tuʔu *stand* | > | *túú |
| *taʔu *season* | > | *taú |
| *leʔo *speech* | > | *lewe *tongue, speech* |
| *paʔoRu *new* | > | *fau |

(There seems to be one exception to this pattern, however, as POC *ʔo(n)ta *raw* is reflected in PTK as *kocaa *raw food*. The rest of MC is ambivalent regarding the *k in this item, however, as *k is not reflected in KIR or KSR. It may well be that the *k in TK and PP is a reflex of a fossilised causative prefix *ka-.)[17]

With only one apparent exception, POC *ŋ is reflected in PTK as *ŋ. The exception is the word for *hermit crab*, which has been reconstructed for POC by Lincoln (ms) as *uŋa, apparently on the basis of FIJ and PPN. The PTK form for *hermit crab* is *umwa.

| POC *ŋ | > | PTK *ŋ | |
|---|---|---|---|
| *aŋin *wind* | > | *aŋi | |
| *(n)toŋo *mangrove* | > | *toŋo | |
| *laŋo *fly* | > | *laŋo | |
| *taŋi(s) *weep* | > | *taŋi | |
| *ŋuu *hum, speak low* | > | *ŋúú | |
| *taliŋa *ear* | > | *taliŋa | |
| *ŋapa *fathom* | > | *ŋafa | |
| *siŋi *push, thrust* | > | *θiiŋi | |
| *saŋa *thigh, crotch* | > | *θaŋa | *inside of upper thigh* |
| *deŋa *turmeric* | > | *raŋa | |
| *taŋiRi *kind of fish* | > | *taŋiri | *yellow fin tuna* |

## 3.3. The coronal series of consonants

With only a few exceptions, POC *t is retained as PTK *t:

| POC *t | > | PTK *t | |
|---|---|---|---|
| *ta *one* | > | *ta-/*te- | |
| *tau *man, person* | > | *tau- | prefix to clan names |
| *ʔatop *thatch* | > | *ato | |
| *mata *eye* | > | *mata | |
| *mate *die* | > | *mate | |
| *tama *father* | > | *tama | |
| *tano *earth* | > | *tano | |
| *toko(n) *staff, pole* | > | *toko | |
| *ata(s) *on top* | > | *ata | |
| *tasi(m) *sharpen* | > | *taim- | |
| *tane *skin disease* | > | *tani | |
| *tusi,u(k) *point* | > | *tiθi | |
| *mata *raw, new* | > | *(a)mata | |
| *loto *boil, abscess* | > | *loto | |
| *pitu *seven* | > | *fitu | |
| *tansik *seawater* | > | *taθi | |
| *tido *look at, gaze* | > | *tiro | |
| *Ratu *one hundred* | > | *(-ŋa)ratu | classifier for thousands |

The exceptions are of several types: (1) POC *watu *thither, toward addressee* and *natu *offspring*, in both of which the *t is lost in all TK languages.[18] (2) The *t in POC *pati *four* is lost in TK and, in fact, in all MC languages. (3) POC *t is apparently reflected as PTK *θ in *kato *basket* > *kaθo *basket*, *tale *return, circumscribe* > *θale *walk around in circles, wander*, and, perhaps, *kita *see* > *kuθθa *find, look for*. These may be borrowings, but if so the sources are not known. (4) The TK forms for *urine* appear to reflect POC *tid,Ri *spurt, urine*, but the reconstruction is problematical. Adding to the problem is the fact that PTK reflects POC *sidi(t) *semen, masturbation* as *tiri *masturbate*. (The expected reflex of *sidi(t) is *θiri.) The TK data-sets for the two POC forms are as follows:

| | | |
|---|---|---|
| POC | *tid,Ri *spurt, urine* | *sidi(t) *semen, masturbation* |
| PTK | *tiri(?) *urine* | *tiri *masturbate* |
| TRK | siir | ir |
| ULI | (not attested) | sir |
| PUA | (not attested) | (not attested) |
| MRT | siir | ir |
| PUL | hiir | ir |
| STW | siir | ir |
| CRL | siir | ir |
| WOL | (not attested) | siri |

The TK forms reflecting POC *sidi(t) correspond exactly, but unexpectedly, with a PTK *tiri, while those that appear to be cognate with POC *tid,Ri are aberrant in that the languages attesting the form fail to lose *t before the high vowel. The ULI and WOL forms for *urine*, where we might on the evidence expect a word homophonous with that for *masturbate*, is the euphemism kkaleoleo, literally *to make puddles*. One possible explanation of the aberrant reflexes for *urine* in the other languages is that the loss of *t would have led to homophony with the terms for *masturbate*, but this 'explanation' requires us in turn to postulate that the two forms were homophones in PTK. Presumably, the solution lies elsewhere.[19]

POC *nt merges with POC *nd as PTK *c, with no apparent exceptions:

| | | | |
|---|---|---|---|
| POC | *punti *banana* | > PTK | *úcú |
| | *ki(n)ta 1st pl inc pron | > | *kica |
| | *-(n)ta 1st pl incl poss pron | > | *-ca |
| | *?o(n)ta *raw* | > | *kocaa |
| | *panda(n) *pandanus* | > | *faca |
| | *ndaŋma *forehead* | > | *camwa |
| | *(n)danum *water* | > | *canu |
| | *(n)dau *leaf* | > | *cau |
| | *nduRi *bone*[20] | > | *cúú |
| | *(n)tau *person* (?)[21] | > | *-caú anim num cls |

This merger also occurs in the other MC languages, in one language of the Admiralties, and in many OC languages of the New Guinea North Coast (Ross 1977) but not, to my knowledge, elsewhere in OC.

With some exceptions (see below) POC *s, *ns and *j merge as PTK *θ:

| POC *s | > PTK *θ | |
|---|---|---|
| *masaki *pain* | > | *maθaki *ill* |
| *kaso *rafter* | > | *kaθo |
| *koso *husk coconuts* | > | *koθo |
| *masawa *open sea* | > | *maθawa |
| *?asu *smoke* | > | *aθu |
| *sau(?) *outside, far off* | > | *θaau[22] |
| *?aRusa *current* | > | *auθa |
| *saRe *tear* | > | *θaari(ŋ-)[23] |
| *isu *nose* | > | *(pwa)uθu |
| *siŋi *push, thrust* | > | *θiiŋi[24] |
| *sedu *hiccough* | > | *(ma)θaru |
| *?asu *gall bladder* | > | *aθi |

POC *ns > PTK *θ

| | | |
|---|---|---|
| *anunsa *small island* | > | *anúθa |
| *nsama *outrigger* | > | *θama |
| *(n)sa *one* | > | *(te-)θa |
| *mansu(r,R) *full, abundant* | > | *maθu |
| *nunsi,o *squid* | > | *ŋúθi/*ŋiθo[25] |
| *ʔunsa(n) *rain* | > | *uθa |
| *tansi(k) *sea water* | > | *taθi |
| *wa(n)se *count* | > | *waθe |
| *(n)siwa *nine* | > | *θiwa |
| *(n)su(n)su *breast* | > | *θuθu[26] |
| *(n)saŋa *thigh, crotch* | > | *θaŋa *inside of upper thighs* |
| *(n)soko *arrive, come* | > | *θoko |

POC *j > PTK *θ

| | | |
|---|---|---|
| *aja(n) *name* | > | *iθa[27] |
| *maja *dry, low tide* | > | *(m)maθa |
| *pija *how much, how many* | > | *fiθa |
| *Ruja(n) *load, cargo* | > | *uθa |
| *(ŋ)ija *when?* | > | *(i-)n,ŋaiθa[28] |

The exceptions are of two types: (1) unexpected loss of *s/*ns in: *masi *breadfruit* > *mai,[29] *(n)soko *caught, captured* > *(y)oko, and *monse *sleep* > *(mw)mwee; and (2) *s/*ns reflected as PTK *t in: *so(ŋ)ka(r) *cross-seat in a canoe* > *toa, *suli *shoot, sucker* > *tili, *sidi(t) *semen, masturbation* > *tiri *masturbate,* *nsiŋi *fart* > *tiŋi,[30] and *(n)sai *who?* > (i-)te,au.[31] So far, these exceptions defy explanation.

Very unexpectedly, TK provides external evidence for the segment *nj, which has been reconstructed for several Siassi (SIA) subgroups in northern New Guinea (Ross 1977). Although only four of the forms reconstructed with *nj by Ross are reflected in PTK, in each instance the segment corresponding to the *nj is lost: *njalan *path, road* > PTK *ala; *kianjo *outrigger boom* > kiao; *tanjim *sharpen* > *taim-; and *panji[32] *arm, wing* > *paú.[33] Ross writes that the SIA distinction between *ns and *nj "is a feature found nowhere else in Oceania" (p.54), and proposes that the distinction resulted from a split of POC *ns in several SIA groups (p.60). While these TK (and other MC[34]) forms may be only coincidental aberrant reflexes of *ns, it appears more likely that they are systematic.

The PTK reflexes of *n and *ñ are only problematical in a few forms before high vowels, where they appear as *ŋ (see below). Otherwise, *n and *ñ merge as PTK *n:

POC *n > PTK *n

| | | |
|---|---|---|
| *niuR *coconut* | > | *núú |
| *ni genitive marker | > | *ni |
| *onom *six* | > | *ono |
| *inu(m) *drink* | > | *únúm- |
| *nana(ʔ) *pus* | > | *nana |
| *ena *away from speaker* | > | *(ke)ena[35] |
| *napo *wave* | > | *nao |
| *natu *offspring* | > | *naú |
| *ʔuna *fish-scale, feather* | > | *úna |

| | | |
|---|---|---|
| *(n)dani(ʔ) *day* | > | *raani[36] |
| *tina *mother* | > | *tina |
| *manaŋ *supernatural power* | > | *mana |
| *eno *lie down* | > | *(w)ono |

POC *ñ > PTK *n

| | | |
|---|---|---|
| *ñamu(k) *mosquito* | > | *namwu |
| *ña 3rd sg poss pron | > | *-na |
| *ñamu *taste, flavour* | > | *(n)naú |
| *maʔañu *float* | > | *maanú |
| *ñoRa *yesterday* | > | *(na)newa[37] |
| *mañawa *breathe, belly* | > | *manawa *life, existence* |
| *kiñit *pinch, pluck* | > | *kini |
| *n,ñopo *stay, dwell, sit* | > | *noo |
| *ʔan,ñitu *ghost, spirit* | > | *anútú |
| *n,ñamo *lagoon* | > | *namwo |

In the environment /a_i, however, OC *n usually becomes the velar nasal *ŋ in PTK:[38]

| | | | |
|---|---|---|---|
| POC | *(n)tani *source* | > PTK | *taŋi |
| | *kani *eat* | > | *kaŋi |
| | *mani *remember* | > | *maŋi *think, remember* |
| | *kani *sharp*[39] | > | *(k)kaŋi |

PTK *ŋ may also reflect POC *n in *nipon *tooth* > *ŋii and *nau 1st sg pron > *ŋaŋú. (In this example, it is the initial velar nasal that shows an aberrant TK reflex of POC *n (also found in other MC languages and in Rotuman). The appearance of the second velar nasal is problematic.[40]) POC *ñ also is reflected as velar *ŋ in PTK *ŋúθi/*ŋiθo *squid* < POC *n,ñunsi,o and *ŋenú *shadow, ghost* < POC *n,ñun,ñu *shadow*, and in TRK, PUL, STW, CRL, WOL wooŋ(i) *turtle* < POC *poñu.[41]

## 3.4. The liquids

While the TK reflexes of POC *l are straighforward, those for *r and, to even greater extent, those for *l are not.

With only two exceptions, POC *l is reflected as PTK *l:

| | | | |
|---|---|---|---|
| POC | *tolu *three* | > PTK | *telu |
| | *lima *five* | > | *lima |
| | *la(m)pa(s) *big* | > | *lapa |
| | *lawa *stuck in throat* | > | *lawa |
| | *lumi *fold, crease* | > | *lúmi |
| | *malu *calm, quiet* | > | *malua *calm, of the sea* |
| | *kali *dig* | > | *keli |
| | *kila *know* | > | *kúla *know, see* |
| | *laŋo *canoe roller* | > | *laŋo |
| | *walu *eight* | > | *walu |
| | *lumpa *hole*[42] | > | *lipwa |
| | *d,lalo *inside, deep* | > | *lalo- *inside* |
| | *s,jola *mix* | > | *θola |
| | *lako *go* | > | *lako dir suff: *away, go* |
| PAN | *ʔalejaw *day*[43] | > | *alo *sun*[44] |

In the two exceptional cases, POC *l is replaced by PTK *r:

| | | |
|---|---|---|
| *kula *circumcise* | > | *kura |
| *sakulaya *sword-fish* | > | *θakul,raara[45] |

POC *d is normally reflected as PTK *r, but there are four instances where the nasal grade *nd is reflected instead (as PTK *c).

POC *d > PTK *r

| | | |
|---|---|---|
| *ŋodo *snore* | > | *ŋoro |
| *ida 3rd pl foc pron | > | *ira |
| *dua *two* | > | *rua |
| *dau *hundred* | > | *(-ŋa-)raú *high number* |
| *dama *light, shine* | > | *(ma-)rama *moon* |
| *(n)doŋo *hear* | > | *roŋo |
| *mudi *behind* | > | *mwuri- |
| *daku *scratch, scrape* | > | *raku |
| *k,ta-dawa *blue/green* | > | *karawa |
| *ʔuda(ŋ) *lobster, shrimp* | > | *úra |
| *data *level, smooth* | > | *rata |
| *do(n)do *dark* | > | *roco |
| *era existential verb[46] | > | *ora |
| *dede *tremble with fear* | > | *rere |

POC *d > PTK *c

| | | |
|---|---|---|
| *d,liki *small, little* | > | *ciki |
| *mada *fermented, soft, ripe* | > | *maca[47] |
| *daRa *blood* | > | *(c)caa |
| *dede *tremble with fear* | > | *cece[48] |

TK (and MC) reflexes of *R are extremely problematical. In the slight majority of cases, *R is lost, while in all but one of the other cases, it merges with *d as PTK *r. The one instance where it seemingly fails to follow either of these patterns is the reflex for POC *tarae *adze*, which is PTK *talee, perhaps indicating a loan. All TK reflexes of *R are presented here:

POC *R > PTK *Ø

| | | |
|---|---|---|
| *maRa *ashamed* | > | *maa |
| *daRa *blood* | > | *(c)caa |
| *paʔoRu *new* | > | *fau |
| *RuʔaRa *neck* | > | *ua |
| *suRi *bone* | > | *cúú |
| *uRa *sinew, vein* | > | *wa(ka) (?)[49] |
| *Ruja *load, cargo* | > | *uθa |
| *nsakaRu *reef* | > | *θakau |
| *kuRita *octopus* | > | *kúúta |
| *noRa *yesterday* | > | *(na)newa |
| *tapud,Ri *conch* | > | *tawii |
| *ʔaRusa *current* | > | *auθa |
| *paRi *ray-fish, skate* | > | *faii[50] |
| *Rapi *evening* | > | *faka-afi *evening, evening meal* |
| *paRu *hibiscus* | > | *fau |
| *Ruŋma *house* | > | *inwa |

POC *R > PTK *r

| | | |
|---|---|---|
| *Ratu *hundred* | > | *(-ŋa-)ratu *thousand* |
| *saRe *tear* | > | *θaari(ŋ-) |
| *puRi *wash* | > | *uro *wash (face)* (?) |
| *ʔapaRa *shoulder* | > | *afara |
| *maRa *light in weight* | > | *mara *fast (e.g. of canoes)* |
| *wakaRa *root* | > | *wakara |
| *paRata *north-west monsoon* | > | *parata *windstorm* |
| *takuRu *back* | > | *takuru |
| *(n)taŋiRi *kind of fish* | > | *taŋiRi *yellow fin tuna* |

George Grace (p.c.) has suggested that *R may tend to be retained in MC between identical vowels, and lost between unlike vowels. Although this suggestion seems generally valid, the first two forms in the above list indicating loss of *R, and the first three forms in the list indicating its retention are counter-examples.

## 3.5. The glides

POC *w is regularly retained in PTK, while *y seems to be lost, with two exceptions.

POC *w > PTK *w

| | | |
|---|---|---|
| *watu *thither* | > | *wau |
| *walu *eight* | > | *walu |
| *awaŋ *mouth* | > | *awa |
| *mawap *yawn* | > | *mawo[51] |
| *pa(ŋ)kiwak *shark* | > | *pakewa |
| *sawaŋ *channel* | > | *θawa |
| *wa(n)se *count, divide* | > | *waθe |
| *kawakawa *kind of fish* | > | *kawakawa *kind of reef fish* |
| *waŋka *canoe* | > | *waa |
| *(n)siwa *nine* | > | *θiwa |
| *mañawa *belly, existence* | > | *manawa |

POC *y > PTK *Ø

| | | |
|---|---|---|
| *mu(yu) 2nd pl poss pron | > | *mii |
| *yaŋo *yellow, ginger* | > | *a o |
| *yapo *fish-line* | > | *ao |
| *mpaya *bait* | > | *paa |

The apparent exceptional reflexes of *y are POC *yadu *mother of pearl* > *taru (?), and *sakulaya *swordfish* > *θakul,raara. The second example seems to be either a case of copying of the preceeding consonant or, perhaps, the accretion of an unknown morpheme *-ra onto a cognate stem (i.e. POC *sakulaya > * θakulaa-ra, where the *y is lost between the two final vowels). The first example is inexplicable if indeed it is cognate.

### 3.6. The vowels

As the above cognate lists indicate, PTK reflexes of POC vowels are generally identical except for the allophony among the high vowels (discussed briefly in Section 2). As most unusual reflexes in the consonant lists have been treated in notes, no further discussion will be provided here.

## 4. CONCLUSIONS

The patterns of regular PTK reflexes of POC presented in Section 3 provide strong evidence for the integrity of the TK subgroup. To my knowledge, there is no other language or language group that combines all of the following phonological innovations (see Jackson, in preparation, for lexical and grammatical innovations of TK):

(1) Loss of POC *p before round vowels;
(2) Loss of POC *ŋk in all environments;
(3) Loss of POC *ʔ in all environments;
(4) Merger of POC *n with *ŋ in the environment /a_i;
(5) Merger of POC *n and *ñ elsewhere;
(6) Merger of POC *s, *ns, and *j;
(7) Separate reflex of POC? *nj;
(8) Merger of POC *nt and *nd;
(9) The pattern of loss of *R and its merger with *r described in Section 3.5 above;
(10) Loss of POC *y.

Innovations (1), (3), (8), and (9) are also attested in other MC languages, and (6) is attested in all MC except KSR (where POC *ns is apparently lost). This fact, in turn, provides some evidence for the integrity of that group. Additional evidence for MC is provided by the sharing of the rule by which POC *mp and *m are merged with *ŋp and *ŋm before round vowels, and by a general MC consistency in reflecting nasal and oral grade segments. (For example, for POC *daRa *blood*, *diki *small*, *mada *ripe*, *(n)dau *leaf*, *(n)danum *water*, *peka *faeces*, *(n)tupa *dērris fish poisōn*, and many ōthers, all MC reflexes of the underlined segments are of nasal grade consonants.) It is beyond the scope of this paper to provide further evidence for a MC grouping (further evidence will be found in Harrison and Jackson, in preparation), but the above observations are suggestive.

Another task that is beyond this paper's scope is to attempt to determine the closest relatives of TK outside MC. However, it is perhaps appropriate at this time to mention that the distinct TK reflexes of *nj discussed in Section 3.3 are not the only characteristics that appear to link TK with languages of the New Guinea north coast. Jackson (in preparation) provides a list of such similarities.

## NOTES

1. I am indebted to a great many people for comments, suggestions, and observations at different times in the development of this paper. Foremost among them are Paul Geraghty, Shelly Harrison, Jeff Marck, Ken Rehg, Ho-min Sohn, Hiroshi Sugita, Jude Wang, and the following Micronesians, who were always willing to answer still more questions about their languages: Marciana Akasy, Jesus Elameto, Rioichy Johnny, Tony Kaipat, Rosario Ngirbabul, Mike Olap, Frank Olopai, Tony Otto, Ancheres Rechim, Tony Tawerilmang, and Joe Tiucheimal. Naturally, however, any errors of reporting or interpretation are my own.

2. Non-Trukic languages are spoken on ten islands in the Carolines: Ponape, Ngatik, Pingelap, Mokil, Kosrae, Nukuoru, and Kapingamarangi in the east, and Yap, Babelthuap and Peleliu in the west. The high islands of Yap and Babelthuap, however, do have populations of Trukic immigrants from neighbouring atolls, and large numbers of Trukic speakers have resettled on Ponape in this century.

3. Goodenough and Sugita (1980) have also identified as Trukic the language spoken by earlier inhabitants of the island of Mapia, approximately 100 miles north of West Irian. They include no supporting evidence, however, and the small amount of data available to me are not conclusive, although Old Mapian does appear 'Micronesian'. (See also Bender 1975.)

4. In addition, I have also referred to my own notes on Saipan Carolinian (CRL), which is very closely related to STW.

5. Nauruan may also be a member of MC; it is difficult to know from the available data. Problems regarding the status of KSR are due in large part to its inconsistent reflexes of POC, probably suggestive of several layers of borrowing. (See Lee and Wang, this volume, and Harrison and Jackson, in preparation.)

6. In fact, pre-MC *t (> PTK *t) was probably palatalised as well, and there is evidence in MC to suggest that it had already developed a sibilant allophone, at least before *i. (See Harrison and Jackson, in preparation.)

7. The glide in this PTK form is expected. The most probable historical stages in pre-TK were: POC *(n)sipo > *θio, > PTK *θiwo.

8. This form is tentatively reconstructed on the basis of Lakalai (LAK) potu *outside*, otu *to go out*, Ponapean (PNP) iei *out*, and the TK forms.

9. The unexpected extra vowel in PTK *faifine must be reconstructed to account for, e.g. TRK feefin and WOL faifilę *woman*. Goodenough (p.c.) has suggested that the initial element of the form may be related to PAN *bahi *female, woman*.

10. The consonant metathesis that occurred in pre-TK to POC *palisi *grass* and *palusa *paddle* is also witnessed in PP, and the metathesised form for 'paddle' is witnessed in KIR, as well, implying that this change may have occurred in pre-MC. (No other MC language attests the forms at all.)

11. The final vowel is unexplained.

12. PTK *pwuu suggests a pre-TK *mpupu.

13. The change from final *u to *i in this form must presumably have antedated the rule velarising *mp before a back round vowel. Otherwise, the PTK form should be *kapwu.

14. TRK reflects only *kami, while the other TK languages reflect *kamami. ULI reflects both forms, although the reflex of *kamami is apparently more frequent.

15. This form is tentatively reconstructed on the basis of the TK forms and Gedaged tamol *man, male person,* Nguna and Sesaki (na)tamwoli *human being,* and perhaps Rotuman famori *human being,* the last three of which Pawley (1979a) assigns to a POC *ta-maquidip *living man.* PTK *tamwoolu, however, does not reflect Pawley's reconstruction.

16. TRK also has a form siim *medium sized tridachna,* which is a more likely reflex of POC *kima. However, no comparable form appears in other TK languages.

17. This analysis is weakened by the fact that the vowel after the *k is short. If the form in fact were derived from a hypothetical *ka-oca-a, then two moras would be expected in the first syllable. However, note that the final syllable in PTK *kocaa (cf. CRL oschaal *his raw food*) is long, which would be expected if the form were a fossilised causative.

18. The *t is reflected in the cognate forms in MRS and KIR, which means that it was certainly present in PMC. The loss of *t before non-low vowels has been spreading at differing rates through the lexicons of all TK languages (see Jackson, in preparation), so it is possible either that the loss in these forms had already occurred by the PTK stage, or that the *t was present in PTK but has since been lost in all TK languages.

19. Dyen (1949) has pointed out that TRK siir is a reflex of PAN *cirit *spray out, urine.* If PAN *c could be shown to be reflected distinctively in other forms in TK, then the mystery of this form's development would be solved, and another segment would be reconstructed for PTK (and POC). However, there is no corroborating evidence, and, as Blust (1978) points out, it is inappropriate to reconstruct a distinction in POC on the basis of one form in one group of languages.

20. POC *nduRi *bone* is also reflected in the Admiralties (Blust 1978) and in Proto-Sepik-Madang of the New Guinea north coast (Ross 1977).

21. Although PTK reflects POC *tau *person* as *tau- *prefix to clan names,* there is also found in MRT, PUL, STW, and CRL the animate number classifier *-caú. Although the latter form may be an innovation, I have tentatively concluded that it derives from a nasal grade alternant of the POC form. Other evidence exists in TK for such alternants (e.g. PTK *fici *to shoot a weapon, such as a slingshot or bow, to spear*; *pici *to snap, spring* (cf. FIJ vidi *to snap*)).

22. The long vowel in PTK *θaau is unexplained.

23. The long vowel is unexplained.

24. The long vowel is unexplained.

25. There is no obvious explanation for the two forms for *squid* in TK. PTK *ŋiθo is reflected in TRK, PUL, WOL, ULI, and PUA; apparent *ŋúθi is reflected in MRT, STW, and CRL.

26. An alternate form *tuθu *breast* must be reconstructed for several TK languages. Interestingly, Ross (1977) has reconstructed the same form for his Proto-North Coast in New Guinea.

27. The high front vowel found in PTK *iθa *name* is unique in MC, but it is witnessed in LAK isa (Chowning 1973).

28. It is possible that the alternate forms for *when?* in TK may reflect a distinction similar to that found in LAK alaisa *when?* (past), gaisa *when?* (future). It is difficult to tell, however, as no single TK language reflects both forms for *when.*

29. Cf. PPN *mei *breadfruit.*

30. TK forms for *fart* are problematical in the same way as those for *urine* discussed above: they reflect failure to lose the *t in TRK, PUL, MRT, STW, and CRL, where such loss is expected.

31. Cf. Gedaged itai *who? (male).*

32. Ross reconstructs this form only for Proto-Manam.

33. The form *pau *arm, wing* is also witnessed in the Admiralties, where there is no other evidence for a distinct *nj. The PTK form, thus, may not in this case reflect loss of that segment. The other three forms remain of interest, however.

34. In addition to TK, MRS and PP reflect PMC *ala *road* and *kiao *outrigger boom* (KIR and KSR reflect an apparent PMC *kiaso for the latter form); all MC languages reflect PMC *pau *arm, wing* and *taim- *sharpen.*

35. This form is probably derived from *ko '2nd sg pron' and *ena *away from the speaker.* It is attested throughout TK, but apparently in no other languages.

36. Vowel length is unexplained.

37. Loss of rounding from a vowel following insertion of a rounded glide is frequent in TK.

38. The TK forms *kaŋi *eat* and *(k)kaŋi *sharp* are also attested elsewhere in MC.

39. This form is tentatively reconstructed on the basis of MC and several languages in the New Hebrides (Tryon 1976).

40. As Harrison (p.c.) has pointed out, the initial *ŋ in PTK *ŋaŋú '1st sg foc pron' may well not reflect POC *n at all, but may instead reflect a prefixal element *ŋ-. Evidence for this is provided by the fact that all KIR focus pronouns reflect such a prefix.

41. MRT, ULI and PUA reflect the expected alveolar nasal.

42. This form is tentatively reconstructed on the basis of TK and PAN *lubaŋ *hole.*

43. This PAN reconstruction is from Blust n.d.

44. Tryon (1976) reports many instances of alo *sun* in the New Hebrides, and Blust (1978) has also reported it for the Admiralty Islands. Presumably, therefore, *ʔalo *sun* must be reconstructed for POC in competition with already reconstructed POC *ʔanso *sun.* PUL, STW, and CRL alet *sun* may reflect the *j of PAN *ʔalejaw, through a hypothetical PTK *aleθi. So far as I know, however, this consonant is not reflected elsewhere in MC.

45. In this form, the expected *l is reflected in all languages except TRK and WOL, and TRK has a doublet with the expected reflex.

46. This form is reconstructed on the basis of PTK *ora, MRS or, and PAN *era 'existential verb'.

47. Reflexes of the nasal grade variants of POC *mada *ripe, fermented, etc.* and *daRa *blood* are attested in the Admiralties (Blust 1978), the North Coast of New Guinea (Ross 1977), and throughout MC.

48. There is apparently no distinction in meaning between PTK *rere and *cece *shake, tremble.*

49. Although a pre-TK *ua could reasonably evolve into PTK *wa (cf. the fact that many CRL speakers say wa and wwa for *fruit* (< PTK *uwa)), I am unable to identify the final *ka in *waka *sinew, vein.*

50. Final vowel length is unexplained.

51. The final vowel in PTK *mawo *yam* is presumably the result of progressive assimilation from the preceding glide. This is unusual, however.

## BIBLIOGRAPHY

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics,* vol.8: *Linguistics in Oceania,* 426-465. The Hague: Mouton.

1975 Mapian: an extinct Micronesian language? Paper presented at the Second Annual Oceanic Linguistics Festival, Linguistic Society of Hawaii, Honolulu.

BENDER, Byron W., et al.

1982 Proto-Micronesian word-list. Computer printout. Department of Linguistics, University of Hawaii.

BIGGS, Bruce

1965 Direct and indirect inheritance in Rotuman. In G.B. Milner and Eugénie A. Henderson, eds *Indo-Pacific linguistic studies,* part 1: *Historical linguistics,* 383-415. Amsterdam: North-Holland. (=*Lingua* 14:383-415.)

BLUST, Robert A.

1972 Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list. *Working Papers in Linguistics, University of Hawaii* 4/1:1-43.

1978 *The Proto-Oceanic palatals.* Memoir No. 43. Wellington: The Polynesian Society.

n.d. Variation in retention rate among Austronesian languages. MS.

CAPELL, Arthur

n.d. (1951) *Grammar and vocabulary of the language of Sonsorol-Tobi*. CIMA Report 30. (Mimeo). Washington: Pacific Science Board. (Republished, 1969, as *Oceania Linguistic Monograph* 12.)

CHOWNING, Ann

1973 Milke's 'New Guinea cluster': the evidence from northwest New Britain. *Oceanic Linguistics* 12:189-244.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

1965 *A sketch of Trukese grammar*. Essay 4. New Haven: American Oriental Society.

ELBERT, Samuel H.

1947 Ulithi-English and English-Ulithi word list. Mimeo. United States Naval Military Government.

1972 *Puluwat dictionary*. *Pacific Linguistics*, C-24.

1974 *Puluwat grammar*. *Pacific Linguistics*, B-29.

GERAGHTY, Paul A.

1979 Topics in Fijian language history. Doctoral dissertation, University of Hawaii, Honolulu. (to appear in *Oceanic Linguistics* Special Publications).

GOODENOUGH, Ward H. and Hiroshi SUGITA

1980 *Trukese-English dictionary*. Memoir 141 of the American Philosophical Society. Philadelphia: American Philosophical Society.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

GRACE, George W, James A. THARP, Peter C. LINCOLN, and Janet BLACK

1980 Proto-Oceanic finderlist. Computer printout. 88pp. [An earlier version appeared as Grace 1969.]

HARRISON, Sheldon P. and Frederick H. JACKSON

in prep- aration The evidence for a Micronesian grouping.

JACKSON, Frederick H.

1978 The synchronic status of the historical rule of final vowel deletion in Saipan Carolinian. Paper presented at the Historical Linguistics Festival of the Linguistic Society of Hawaii. Honolulu.

in prep- aration Aspects of the historical development of the Trukic languages of Micronesia. Doctoral dissertation, University of Hawaii, Honolulu.

JACKSON, Frederick H. et al.

forth- coming *Carolinian-English dictionary*. Honolulu: University Press of Hawaii.

LEE, Kee-dong and Judith W. WANG

1983 Kosraean reflexes of Proto-Oceanic phonemes. See pp. 403-442 in this volume.

LEVY, Richard

1979 Languages of the southeast Solomon Islands and the reconstruction of Proto-Eastern-Oceanic. In Paz Buenaventura Naylor, ed *Austronesian Studies: papers from the Second Eastern Conference on Austronesian Languages,* 213-225. Michigan Papers on South and Southeast Asia, No. 15. Ann Arbor:University of Michigan.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

n.d. Thoughts on Proto-Trukic. MS.

ODA, Sachiko

1977 The syntax of Pulo Annian: a nuclear Micronesian language. Doctoral dissertation, University of Hawaii, Honolulu.

PAWLEY, Andrew K.

1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Readings in Oceanic culture history,* vol.3: 1-142. Pacific Anthropological Records 13. Honolulu: Bernice Pauahi Bishop Museum.

1977 On redefining 'Eastern Oceanic'. MS.

1979a Proto-Oceanic terms for people: a problem in semantic reconstruction. MS.

1979b *New evidence on the position of Rotuman.* Working Papers in Anthropology, Archaeology, Linguistics, Maori Studies No.56. Department of Anthropology, University of Auckland.

n.d. A Proto-Oceanic grammar. MS.

QUACKENBUSH, Edward M.

1968 From Sonsorol to Truk: a dialect chain. Doctoral dissertation, University of Michigan, Ann Arbor. (University Microfilms, 69-2375.)

REHG, Kenneth L.

1983 The origins of 'Compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See pp.53-59 this volume.

ROSS, Malcolm

1977 Relationships of the Austronesian languages of the Sepik and western Madang coast of New Guinea. Mimeo.

SOHN, Ho-min

1975 *Woleaian reference grammar*. Honolulu: University Press of Hawaii.

SOHN, Ho-min and Byron W. BENDER

1973 *A Ulithian grammar. Pacific Linguistics*, C-27.

SOHN, Ho-min, and Anthony F. TAWERILMANG

1976 *Woleaian-English dictionary*. Honolulu: University Press of Hawaii.

SOHN, Ho-min, Anthony TAWERILMANG, Isaac LANGAL, and Celestine YANGILMAU

1977 Consonant shifts and subgrouping in the Sonsorol-Ulithi-Woleai chain. *Language Sciences* 44:19-24.

SUGITA, Hiroshi

in prep- Trukese reference grammar.
aration

TRYON, Darrell T.

1976 *New Hebrides languages: an internal classification. Pacific Linguistics* C-50.

# ON THE HISTORY OF PONAPEAN PHONOLOGY

Kenneth Rehg

This paper[1] presents some preliminary observations on the history of Ponapean (PNP) phonology. First, to provide a context for this discussion, synchronic statements about Ponapean dialectology and phonology are presented. Then, diachronic hypotheses are developed by examining the phenomenon of segmental erosion from the right and by detailing Proto-Oceanic (POC) sources for the phonemes of the Main dialect. Finally, concluding comments are offered along with a list of POC/PNP cognates.

## 1. ABOUT PONAPEAN

PNP is a nuclear Micronesian (MC) language which serves as the lingua-franca of the Ponape district of the United States Trust Territory of the Pacific Islands. At present, PNP is spoken as a first language by approximately 16,000 inhabitants of the islands of Ponape, Ant, and Ngatik.[2] Although no variation studies have been conducted on this language, at least three geographically defined dialects may be easily discerned.

### 1.1. The Main dialect

On the island of Ponape, in the municipalities of Madolenihmw, Uh, Net, and Sokehs, and in the administrative centre of Kolonia, the northern or Main dialect is spoken. If one may speak of a prestige dialect of PNP, it is this one. It is the dialect most widely spoken and its pronunciation has been adopted as the basis for orthographic reform.[3] It is the dialect to be considered in this paper.

### 1.2. The Kiti dialect

On the southern part of the island of Ponape, principally in the municipality of Kiti, the Kiti dialect is spoken. A few Kiti speakers also reside on the atoll of Ant. The Kiti dialect differs from the Main most noticeably in terms of vowel distribution. Many Main dialect forms containing [ɛ] have Kiti cognates with [ɔ] or [a]. Further, a surface contrast between [e] and [ɛ] exists in the Main dialect that is not found in Kiti speech. In an essentially taxonomic analysis of the language, Garvin reports that

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 281-316. *Pacific Linguistics*, C-80, 1984.

> in the Kiti dialect, the only mid front vowel phoneme has a closed allophone [e] before y and w, an open allophone [ɛ] in all other positions; long e· is always as [ɛ·]....[4]

In Section 2.3., it will be suggested that within a generative framework, a very similar statement may be made to account for occurrences of [e] in the Main dialect, and that the surface contrast between [e] and [ɛ] is not phonemic.

## 1.3. Ngatikese

Ngatikese (NGK) is spoken on the atoll of Ngatik, located approximately 90 miles to the south of Ponape. Little information is available concerning this dialect, but other speakers of PNP characterise it as being an English/Ponapean creole. All dialects of PNP contain a large number of English loans, but apparently NGK has borrowed entire phrases from English and employs loans not present in any other PNP dialects. Purportedly this came about as a consequence of contact with New England whalers who, during the mid-nineteenth century, invaded the island, slaughtered the male population, and intermarried with the women. At least one unique phonological feature of NGK has also been observed; that is, the trilled r of other dialects is reflected here as a voiceless velar fricative.

## 2. PONAPEAN PHONOLOGY TODAY

Comparisons between PNP and Proto-Pceanic (POC) are based on the following inventory of systematic phonemes for the Main dialect of Ponapean.[5]

### 2.1. Consonants

| | Bilabial | Apico-Dental | Apico-Alveolar | Retroflex | Dorso-Velar |
|---|---|---|---|---|---|
| Stops | | | | | |
| Plain | p | t | | | k |
| Velarised | $p^w$ | | | | |
| Affricates | | | | ʦ | |
| Fricatives | | | s | | |
| Nasals | | | | | |
| Plain | m | n | | | ŋ |
| Velarised | $m^w$ | | | | |
| Laterals | | l | | | |
| Trills | | | r̃ | | |

For orthographic convenience, these phonemes will subsequently be written as follows:

| The Phoneme | Will be Written |
|---|---|
| $p^w$ | pw |
| $m^w$ | mw |
| $\tilde{r}$ | r |
| ts | t |
| ʻʻ | ʻ |
| ŋ | ng |

All other phonemes will be represented as they are on the preceding chart.

## 2.2. Semi-consonants

For the purpose of this report, two semi-consonants, /y/ and /w/ are postulated. The phonemic status of /y/, however, is uncertain.

## 2.3. Vowels

| | Front Unrounded | Back Rounded |
|---|---|---|
| High | i | u |
| Mid | e | o |
| Low | a | ɔ |

Each of these vowels occurs long as well as short, with length being phonemic. Long vowels in this report will be written doubled.

Garvin, however, in his analysis states:

> The Main dialect of [PNP] has 7 determinate syllabic vowels: a, ɔ, ɛ, o, e, u, i; one indeterminate syllabic vowel: ə; ...all the determinate syllabic vowels have the additional prosodic feature of quantity; ....[6]

Thus, the analysis suggested in this paper differs from Garvin's in two respects. First, no intermediate vowel schwa is postulated. Second, the contrast between e and ɛ is not considered phonemic. Since this position has not before been formally presented, its bases are outlined below.

Note that in Garvin's analysis only determinate vowels may occur long. /ə/, an indeterminate vowel, is always short. It will be argued in this paper, however, that all vowel phonemes occur both long and short, and, when they are short, they all have central allophones conditioned by adjacent consonants.

Consonants in Ponapean form two series — a front series and a back series.

| Front | Back |
|---|---|
| p | pw |
| m | mw |
| t | tʻ |
| l | r |
| n | ng |
| s | - |
| - | k |

This front/back distinction is based on the observation that consonants of the front series centralise short back vowels, and conversely, consonants of the back series centralise short front vowels. Illustrations of this follow.

| Front Vowels | Between Front C's | | Between Back C's | |
|---|---|---|---|---|
| i | [pɪl] | *also* | [rɨr̥] | *secret* |
| e | [mɛm] | *sweet* | [tə̀ŋ] | *tight* |
| a | [paˆp] | *swim* | [kɑ̀ˆk] | *can* |

| Back Vowels | Between Front C's | | Between Back C's | |
|---|---|---|---|---|
| u | [lʊs] | *jump* | [pʷuŋ] | *correct* |
| o | [pɵs] | *explode* | [tòpʷ] | *lush* |
| ɔ | [pɔs] | *hammer* | [rɔŋ] | *burned* |

The examples above illustrate short vowel phonemes in 'pure' environments; that is, in positions between either front or back consonants. However, when these vowels occur between consonants of different series, as they often do, the vowel quality detectably glides in response to the adjacent consonant, with the following consonant dominating. An example of this is [mᵉək] *rust* where the onset of the vowel is front, but the primary quality of the vowel is central.

Further justification for setting up two series of consonants is the interesting fact that front/back consonants form pairs, as they are listed in the preceding chart, and these pairs, excluding s and k, almost never occur within the same morpheme.[7] Thus, a morpheme like pVp or pwVpw is acceptable, but not pVpw or pwVp. Apparently, there is a morpheme structure rule in PNP which involves consonant harmony of this nature.

Central vowels, then, even within a taxonomic framework, are not phonemic in PNP. Instead, consonants condition centralisation in a manner reminiscent of Marshallese. The other point of disparity with Garvin's analysis, though, the number of mid front vowels, is a consequence of the generativist position taken in this paper as opposed to the structuralist position taken by Garvin.

Garvin postulated two mid front vowels for the Main dialect, /e/ and /ɛ/, based on minimal pairs like [seysey] *to cut hair* and [sɛysɛy] *to paddle*; [eet] *bag* and [ɛɛt] *one*. However, other considerations are relevant here.

1) Excluding loan words, [e] never occurs long except as a consequence of compensatory lengthening. Thus, one finds forms like [ee̊t], but in the construct it is [etin]. [ɛɛt] is also long in the construct, [ɛɛtin].

2) [e] is always followed by a high vowel or glide, either on the surface as in [seysey] or when affixed as in [etin]. [e] never occurs word finally.

3) Where minimal pairs occur between short [e] and [ɛ], the form containing [ɛ] has a related form with the vowel [a]. [sɛysɛy] means *to paddle*, but [saay] is *a paddling trip*.

4) Other occurrences of short [ɛ] not alternating with [a] are not followed by high vowels or glides. Thus, [tɛ] *or* or [mɛt̀] *here*.

These observations then suggest that one postulate a single mid front vowel with a phonetic norm [ɛ] and a raised short allophone [e]. The following informal

rules and base forms suggest how surface minimal pairs can be accounted for. The seemingly unnecessary complexity of the rules is motivated by other facts of the language not discussed here.

| | | |
|---|---|---|
| Base | sɛi | sai |
| Glide Insertion | sɛyi | sayi |
| Short ɛ → e/___C$_o$ [-cons, +high] | seyi | -- |
| Short a → ɛ/___C$_o$ [+syllabic, +high, -back] # | -- | sɛyi |
| Final Vowel Deletion | sey | sɛy |
| Surface Form (Reduplicated) | seysey | sɛysɛy |

To account for nouns like [eet], [ɛɛt], and [saay], it is necessary to add a lengthening rule which applies only to nouns of the canonical form #(C)VC(V)#.

| | | | |
|---|---|---|---|
| Base | ɛti | ɛɛti | sai |
| Glide Insertion | -- | -- | sayi |
| Short ɛ → e/__C$_o$ [-cons, +high] | eti | -- | -- |
| Lengthening | eeti | -- | saayi |
| Short a → ɛ/__C$_o$ [+syllabic, +high, -back] # | -- | -- | -- |
| Final Vowel Deletion | eet | ɛɛt | saay |

Further discussion of the lengthening rule and its peculiar order with respect to final vowel deletion will be offered in Section 3.2.

It is necessary, then, within a generative framework, to establish only a single mid front vowel phoneme. However, since the theoretical position on which this analysis is based is not accepted by all scholars, and so that information may not be lost to them, the subphonemic contrast of [e] and [ɛ] will continue to be indicated in all transcription.

## 3. EROSION FROM THE RIGHT

The deletion of final POC consonants and vowels is a phenomenon characteristic of MC languages. The status of these final segment deletion rules, though, whether synchronic or diachronic, is not easily determined.

### 3.1. Final consonant deletion

In general, final POC consonants are lost in PNP.

| | | |
|---|---|---|
| *alap | alɛ | *take* |
| *?atop | ɔɔs | *thatch* |
| *awang | aaw | *mouth* |
| *ngkinit | kinii | *pinch* |
| *manang | manaman | *spiritual power* |
| *onom | wene- | *six* |
| *pakiwak | pako | *shark* |
| *Ruap | uu | *tide* |
| *?uling | iliil | *steer* |

In five instances, however, final POC consonants are retained.

| | | |
|---|---|---|
| *inu(m) | nim | *drink* |
| *pang(ou)(n) | pangin | *awaken* |
| *piti(k) | piṭik | *spring up* |
| *sakan | sakan | *eat* - honorific |
| *tokon | sokon | *cane* |

The retention of the final *n in sokon is unexplainable. A possible explanation for the first four items, though, is suggested by the fact that they are all transitive verbs, three of which have derived intransitive forms which do not reflect the final consonant.

| Transitive | Intransitive |
|---|---|
| pangin | mpɛng |
| piṭik | piṭ |
| sakan | sak |

The retention of the final POC consonant in these transitive verbs might be explained then by postulating the presence of a transitivising suffix which blocked the application of the final consonant deletion rule. In the intransitive forms, where this suffix was not present, the final consonant dropped. This position seems highly tenable in light of synchronic facts about the language, where, in fact, transitive suffixes do occur.

A problem with this solution is that final consonants were deleted in some transitive forms.

| | | | |
|---|---|---|---|
| *alap | alɛ | tr. | *take* |
| *ngkinit | kinii | tr. | |
| | kinikin | intr. | *pinch* |
| *?uling | iliilii | tr. | |
| | iliil | intr. | *steer* |

Since alɛ occurs only transitive, one might assume that no transitive suffix occurred with this form. In the case of the other two verbs, the intransitive form appears to be the base. Note the presence of the transitivising suffix -ii. Which form of the verb was selected as the base, though, ultimately appears to be morphologically conditioned.

Further complexities with transitive/intransitive pairs are suggested by these verbs:

| | | | |
|---|---|---|---|
| *sa?it | tɛɛk | tr. | |
| | teytey | intr. | *sew* |
| *iriR | irip | tr. | |
| | irir | intr. | *fan* |
| *sampo | tapɛr | tr. | |
| | tapatap | intr. | *catch* |
| *puki | wikit | tr. | |
| | wik | intr. | *turn* |

The transitive form of the verb in each of the above cases exhibits a final consonant. In no case, though, is this final consonant a reflex of the final reconstructed segment of the POC form. What the source is of these final consonants is at this point unexplainable, but, of course, this is a common problem with Oceanic languages.

## 3.2. Final vowel deletion

A comparison of POC reconstructions and PNP cognates illustrates the motivation for a final vowel deletion rule.

| | | |
|---|---|---|
| *api | aay | *fire* |
| *dongo | rong | *hear* |
| *kuku | kiik | *nail* |
| *limu | liimw | *sponge, moss* |
| *masawa | mataw | *sea* |
| *nana | naan | *pus* |
| *puko | uuk | *net* |
| *sala | aal | *road* |
| *solo | tool | *mountain* |
| *tali | saal | *rope* |
| *tuki | suk | *pound* |
| *pitu?u | usu | *star* |

The effect of this rule is to delete a single final vowel mora. Thus, consider *piti?u. Since *? apparently goes to Ø in all MC languages, one may assume the loss of this consonant first. This results in an intermediate form pituu. By then deleting a final vowel mora, one can account for the short final vowel which is present in the PNP form, usu.

Further, it is clear that final vowel deletion must follow final consonant deletion in order to account for forms like these:

| | *?atop *thatch* | *awang *mouth* |
|---|---|---|
| Final C deletion | ?ato | awa |
| Other rules | ɔɔso | aawa |
| Final V deletion | ɔɔs | aaw |

What is not clear is the extent to which this rule is a diachronic phenomenon of the language. As previously discussed in Section 2.3., there is justification for establishing a synchronic final vowel deletion rule. Notice:

| Elicitation Form | Construct Form |
|---|---|
| kiik | kikin |
| laang | lɛngin |
| aat | ɛtɛn |

To derive these forms, the following bases and rules are postulated.

**Bases**

kiki *nail*, langi *sky*, ata *name*, ni 'construct suffix'.[8]

Compensatory Lengthening (CL)

$$\#[\,([+\text{cons}])[+\text{syll}][-\text{syll}]([+\text{syll}])\,]_N\# \Rightarrow \#[\,([+\text{cons}])\begin{bmatrix}+\text{syll}\\+\text{long}\end{bmatrix}[-\text{syll}]([+\text{syll}])\,]_N\#$$

This rule lengthens the first vowel of nouns with the canonical shape described above.

Vowel Raising (VR)

$$\begin{bmatrix}+\text{syllabic}\\-\text{back}\\+\text{low}\\-\text{long}\end{bmatrix} \rightarrow [-\text{low}]/\_\_C_0\begin{bmatrix}+\text{syllabic}\\-\text{back}\\+\text{high}\end{bmatrix}\#$$

This rule raises short a to e (phonetically epsilon) when followed by i in word final position.

Vowel Copying (VC)

$$\begin{bmatrix}+\text{syllabic}\\-\text{back}\\+\text{low}\\-\text{long}\end{bmatrix} \rightarrow [-\text{low}]/\_\_C_1\begin{bmatrix}+\text{syllabic}\\-\text{back}\\-\text{low}\\-\text{long}\end{bmatrix}C_0\begin{bmatrix}+\text{syllabic}\\-\text{back}\\+\text{high}\end{bmatrix}\#$$

This is a vowel harmony rule which raises a to e (phonetically epsilon). This is an approximation of this rule. Its precise nature is still not understood.

Final Vowel Deletion

$$V \rightarrow \emptyset/VC(V)\_\_\#$$

Note that this rule does not delete the final vowel of monosyllables.

Based upon a large body of synchronic data, these rules appear to be well motivated. They are applied to the above base forms as follows:

| | kiki | kiki+ni | langi | langi+ni | ata | ata+ni |
|---|---|---|---|---|---|---|
| CL | kiiki | -- | laangi | -- | aata | -- |
| VR | -- | -- | -- | -- | -- | atɛ+ni |
| VC | -- | -- | -- | lɛngi+ni | -- | ɛtɛ+ni |
| FVD | kiik | kiki+n | laang | lɛngi+n | aat | ɛtɛ+n |
| Surface | kiik | kikin | laang | lɛngin | aat | ɛtɛn |

These rules, in this order, generate the correct PNP surface forms. It would at first seem curious that compensatory lengthening is synchronically ordered prior to final vowel deletion. But, this is very likely the order in which these changes took place diachronically. Some western Trukic languages, Woleaian for example, still retain voiceless vowels in final position. This suggests that final vowels were first devoiced, compensatory lengthening then took place, and finally voiceless vowels were deleted.

A comparison of the base forms that have been postulated for the above nouns with their POC reconstructions illustrates that for at least some forms in the language, POC final vowels have been retained, although their quality may be modified. Thus:

| POC Reconstruction | PNP Base |
|---|---|
| *kuku | kiki |
| *langi(t) | langi |
| *ansa(n) | ata |

The problem is that synchronically it is probably not possible to postulate a final vowel for all bases where POC reconstructions suggest one. Indeed, as one would expect, final vowels appear to be retained only in those forms which take closely bound suffixes. Otherwise, final vowels are lost with subsequent restructuring of the lexicon. Thus, if one is to reconstruct to the level of the phonological base of PNP, a diachronic final vowel deletion is required which will account for this restructuring, and a synchronic final vowel deletion rule is necessary to generate correct surface forms for those items which do retain the final vowel. Unfortunately, too little is known about PNP to establish reliable base forms for all items cognate with POC. The extent of restructuring that has taken place, then, remains unknown. It is consequently necessary in this paper to reconstruct, not to the base, but to an intermediate point of derivation approximately equivalent to the level of the taxonomic phoneme.

## 4. POC SOURCES FOR PNP CONSONANTS

Following is an inventory of POC consonants.[9]

| | p | t | d | s | k | ? |
|---|---|---|---|---|---|---|
| ngp | mp | nt | nd | ns | ngk | |
| ngm | m | n | | ñ | ng | |
| | w | l | r | y | R | |

PNP reflexes of these POC consonants will be presented in the following discussion. The procedure will be to discuss first the laryngeal *?, then the velar consonants, the coronal consonants, and finally the labial consonants. In addition, comments on glide insertion will be offered along with some observations on the source of initial consonant clusters in PNP. Lists of regular consonantal correspondences will not be exhaustive. Lists of exceptions will be.

### 4.1. The laryngeal *?

*? in PNP, as probably in all other MC languages, is reflected as ø. Examples follow:

| | | |
|---|---|---|
| *?ate | ɛɛ | *liver* |
| *?atop | ɔɔs | *thatch* |
| *ma?udi(p) | mowr | *life* |
| *nana(?) | naan | *pus* |
| *pitu?u | usu | *star* |
| *sa?it | teytey | *sew* |
| *?uling | iliil | *steer* |

In the following example, w appears to be a reflex of *?.

| | | |
|---|---|---|
| *?una(p) | winɛ | *fish scales (its)* |

In Section 4.5., however, it will be argued that this w can be accounted for by a glide insertion rule.

## 4.2. Velar consonants

### 4.2.1. *k

*k > k

| | | |
|---|---|---|
| *iku | iik | *tail* |
| *ka(Rdr)angpa | kapw | *new* |
| *kamu | komw | *you* - honorific |
| *kuku | kiik | *nail* |
| *kuli(t) | kiil | *skin* |
| *laki | laak | *penis* |
| *masaki(t) | mɛtɛk | *pain* |
| *puki | wik | *turn* |
| *puko | uuk | *net* |
| *sakan | sakan | *eat* - honorific |

A single irregular reflex of *k occurs: *k > Ø.

| | | |
|---|---|---|
| *nsake | -ta | *upwards* |

### 4.2.2. *ngk

*ngk > k

| | | |
|---|---|---|
| *ngkinit | kinii | *pinch* |
| *sangka(?) | tak | *ride (vehicle)* |
| *mpengka | pwɛɛk | *bat* |

### 4.2.3. *R

*R > Ø in seven cases.

| | | |
|---|---|---|
| *daRa(?) | nta | *blood* |
| *ka(Rdr)angpa | kapw | *new* |
| *kuRita | kiis | *squid* |
| *mauRi | mɛyng | *left* |

| | | |
|---|---|---|
| *Ripa | mpɛ | *close (to it)* |
| *Ruap | uu | *tide* |
| *nsakaRu | tɛkɛ | *island* |

Note that in all these cognates except the first two, *R is deleted before a high vowel. The first two forms, however, are also irregular in other respects. The initial segment in *daRa(?) should be nasal rather than oral grade to account for the retroflex in the PNP cognate. kapw from *ka(Rdr)angpa should reflect a long vowel rather than a short one.

*R > r in two cases.

| | | |
|---|---|---|
| *?apaRa | apɛrɛ | *shoulder (its)* |
| *meRa | mer | *rust* |

Here, *R is preserved before *a. However, these cognates are also suspect. The first is archaic and is retained only in the highest honorific speech. Few speakers of PNP know this word. The second exhibits a semantic shift from POC 'reddish' to PNP 'rust'. Also, the construct form of this word is merin, reflecting the wrong reflex for the POC final vowel.

Thus, two positions may be taken with respect to *R. A claim might be made that *R deletes before high vowels only, with two possible exceptions, or one might argue that *R deletes everywhere also with two possible exceptions. It is impossible at this point to determine which hypothesis is correct.

### 4.2.4. *ng

*ng > ng

| | | |
|---|---|---|
| *angoango | ɔngɔɔng | *yellow* |
| *dongo | rong | *hear* |
| *langi(t) | laang | *sky* |
| *lango | lɔɔng | *fly* |
| *ngapulu(?) | ngowl | *ten* |
| *ngipo | ngii | *tooth* |
| *pang(ou)(n) | pangin | *awaken* |
| *sanga | taang | *thigh* |
| *talinga | salɛng | *ear* |

## 4.3. Coronal consonants

### 4.3.1. *s

*s > t

| | | |
|---|---|---|
| *asang | ɛtɛ | *gill (its)* |
| *masaki(t) | mɛtɛk | *pain* |
| *sa?it | teytey | *sew* |
| *masawa | mataw | *open sea* |
| *sanga | taang | *thigh* |
| *sau(?) | too | *far off* |

| | | |
|---|---|---|
| *solo | tool | *mountain* |
| *sulu | tiil | *torch* |
| *tusu | it | *point* |
| *sampo | tapatap | *catch* |
| *sili | til | *penetrate* |

In two instances, *s > Ø.

| | | |
|---|---|---|
| *sala | aal | *road* |
| *suli | iil | *shoot, sucker* |

In one unexplainable instance, *s > s.

| | | |
|---|---|---|
| *sakan | sakan | *eat* - honorific |

### 4.3.2. *ns

*ns > t

| | | |
|---|---|---|
| *ansa(n) | aat | *name* |
| *nuns(io) | nuut | *squid* |
| *mansu(rR) | mɛt | *full* |
| *pinsiko | utuk | *flesh* |
| *nsama | taam | *outrigger* |
| *nsiwa | tuwa- | *nine* |
| *nsoka | tok | *stab* |
| *tansi(k) | seet | *sea* |

In one case, *ns > Ø.

| | | |
|---|---|---|
| *nsangi | aang | *wind* |

Thus, the regular reflex of both oral and nasal grade *s is t. In three instances, *(n)s > Ø. This sporadic deletion of *(n)s with no obvious phonological conditioning is apparently characteristic of MC languages. It has also been reported for Kusaiean, Mokilese (MOK), Trukese (TRK), and Ulithian.[10]

### 4.3.3. *d

*d > r

| | | |
|---|---|---|
| *-da | -ra- | *their* |
| *da?a(n) | raa | *branch, twig* |
| *dongo | rong | *hear* |
| *dua | ria- | *two* |
| *mada | maar | *preserved breadfruit* |
| *matudu | mɛyr | *sleep* |
| *ma?udi(p) | mowr | *life* |
| *madama | maram | *moon* |

In two instances oral grade *d shows a nasal grade reflex.

| | | |
|---|---|---|
| *daRa(?) | nṭa | *blood* |
| *(dl)iki | ṭik | *small* |

### 4.3.4. *r

*r > r

| | | | |
|---|---|---|---|
| *iriR | irip | trans. | |
| | irir | intr. | *fan* |
| *raku | raki | | *scratch* |

In addition, the following ambiguous cases occur.

| | | |
|---|---|---|
| *(dr)an(i) | raan | *day* |
| *(dr)on(st)o | rot | *dark* |
| *ka(dr)u | karu(pwut) | *scratch* |
| *pi(dr)i | pir | *twist* |

This ambiguity is of no consequence, however, since in PNP both *d and *r are regularly reflected as r.

### 4.3.5. *nd

*nd > t'

| | | |
|---|---|---|
| *ndangma | t'aamw | *forehead* |
| *ndamu | t'amw(ɛ) | *lick* |
| *ndau | t'ɛɛ | *leaf* |

### 4.3.6. *nt

*nt > t'

| | | |
|---|---|---|
| *(dr)on(st)o | rot' | *dark* |
| *kinta | kit'- | 1st pl |
| *munta | mwut' | *pulverised* |
| *pinta | piit' | *split leaves of pandanus* |
| *punti | uut' | *banana* |
| *nta | -ta- | *our* |

Thus, *nd and *nt have fallen together, the normal PNP reflex being t'.

### 4.3.7. *t

*t > s in these cases.

| | | |
|---|---|---|
| *tali | saal | *rope* |
| *talinga | salɛng | *ear* |
| *tama | saam | *father* |
| *tangi(s) | sɛng | *weep* |
| *tolu | sili- | *three* |

| | | |
|---|---|---|
| *topu | seew | *sugar cane* |
| *ʔatop | ɔɔs | *thatch* |
| *mata | maas | *eye* |
| *ngmata | mwaas | *worm* |
| *mputo | pwuus | *navel* |
| *uto | uus | *float of fishing net* |

*t > Ø in the following examples.

| | | |
|---|---|---|
| *tupa | uup | *fish poison* |
| *tusu | it | *point* |
| *tuʔu | uu | *stand* |
| *tumpu | ipw | *parentage* |
| *tina | iin | *mother* |
| *natu | nɛɛ | *poss. for children* |
| *matudu | mɛyr | *sleep* |
| *anitu | ɛni | *ghost* |

The preceding forms illustrate that *t > Ø before high vowels, and that elsewhere *t > s. This s reflex of *t, however, is probably a rather recent development. Luther Gulick, in a generally insightful article published in 1880, entitled "A Vocabulary of the Ponape Dialect",[11] describes no s for PNP. Instead, he states:

> ...there is a series of variable and interchangeable sounds, including our 'g' soft or j, and the corresponding surd ch, and verging sometimes even upon sh and ts, for which the single character j is employed here, as being on the whole, the truest rendering of the entire group.

All forms which he spells with j are in modern PNP pronounced with an s, which is phonetically palatalised. In MOK, a language closely related to PNP, *t is still reflected as a voiceless palatal affricate. Thus, it seems reasonable to assume that in PNP, *t first went to a voiceless palatal affricate, and only within the past one hundred years developed the palatalised s.[12]

In two additional cases *t > Ø.

| | | |
|---|---|---|
| *ʔate | ɛɛ | *liver* |
| *mate | mɛɛ(la) | *die* |

These forms would appear to be exceptions to the rule stated above that *t deletes only before high vowels. It may be, though, that e provides a special environment. There still remain these irregular reflexes of *t.

| | | |
|---|---|---|
| *tuki | suk | *pound* |
| *pitu | isi- | *seven* |
| *pituʔu | usu | *star* |

Here, *t is in an environment before a high vowel, but it has an s rather than a Ø reflex. Other MC languages also exhibit irregular reflexes of *t. Thus, as with the problems of *s, this may be a phenomenon of Proto-Micronesian (PMC) rather than of PNP.

In a single instance, *t shows a nasal rather than the expected oral grade reflex.

| | | |
|---|---|---|
| *piti(k) | pi̱tik | *spring up* |

### 4.3.8. *n

*n > n

| | | |
|---|---|---|
| *anitu | ɛni | *ghost* |
| *inu(m) | nim | *drink* |
| *manang | manaman | *spiritual power* |
| *manu(k) | maan | *bird, animal* |
| *nana(ʔ) | naan | *pus* |
| *niu(R) | nii | *coconut* |
| *pani | paan | *bait* |
| *tina | iin | *mother* |

*n > Ø

| | | |
|---|---|---|
| *ponu | weey | *turtle* |
| *namu(k) | aamw | *bug* |
| *pan(i) | pɛɛ | *wing, arm (its)* |
| *tanum | sowsow | *grave* |

Except for *pan(i), where one might assume *n was in final position, no plausible phonological conditioning may be established to account for the deletion of *n.

### 4.3.9. *l

*l > l

| | | |
|---|---|---|
| *alap | alɛ | *take* |
| *langi(t) | laang | *sky* |
| *lima | lima- | *five* |
| *limu | liimw | *moss, sponge* |
| *pili(ʔ) | pil | *choose* |
| *mpulu(t) | pwiil | *gum* |
| *sala(n) | aal | *road* |
| *solo | tool | *mountain* |
| *tali | saal | *rope* |
| *tolu | sili- | *three* |
| *ʔulunga | ulung or uluul | *pillow* |

This is a regular correspondence. There are no exceptions in the data.

## 4.4. Labial consonants

### 4.4.1. *p

*p > p

| | | |
|---|---|---|
| *pa- | pey | *fight* |
| *pa- | pa- | causative prefix |
| *paʔa | pa(ki) | *cut yams for planting* |
| *paʔa | paa | *leaf* |
| *pada | -par | *pandanus* |
| *pakiwak | pako | *shark* |

| | | |
|---|---|---|
| *pan(i) | pɛɛ | *wing (its)* |
| *pang(ou)(n) | pangin | *awaken* |
| *pat | paa- | *four* |
| *pi(dr)i | pir | *spin, twist* |
| *pili(?) | pil | *choose* |
| *pali | pɛl | *respect* |
| *tupa | uup | *fish poison* |

*p > Ø

| | | |
|---|---|---|
| *punti | uuṭ | *banana* |
| *puko | uuk | *net* |
| *pua(?) | waa | *fruit* |
| *ngapulu(?) | ngowl | *ten* |
| *puki | wik | *turn* |
| *tapu- | sowi | *conch shell trumpet* |
| *topu | seew | *sugar cane* |
| *tapu(n)i | sow | *afterbirth* |
| *ponu | weey | *turtle* |
| *mapo | mo | *heal* |
| *nsipo | -ti | *downwards* |
| *ngipo | ngii | *tooth* |

The PNP cognates reflecting *p listed above suggest that *p is lost before round vowels and is retained elsewhere. In some forms, it appears that *p > w; however, it will be argued in Section 4.5. that the w in these forms occurs as a consequence of glide insertion.

The following two cognates present possible ordering arguments for the deletion of *p.

| | | |
|---|---|---|
| *pinsiko | utuk | *flesh* |
| *pitu?u | usu | *star* |

These forms suggest the existence of a vowel harmony rule, the effect of which was to back and round *i to u in concord with the final round vowel of the morpheme. *p, then in the proper environment, was deleted.

Now note this cognate:

| | | |
|---|---|---|
| *pitu | isi- | *seven* |

This form suggests that the deletion of *p should precede the fronting of final short *u to i, a rule which will be discussed in Section 5.1.2. Thus, PNP cognates of *pinsiko, *pitu?u, and *pitu may have evolved as a consequence of these ordered rules.

| POC Forms | *pinsiko | *pitu?u | *pitu |
|---|---|---|---|
| Other Rules | pitiko | pisuu | pisu |
| Vowel Harmony | putuko | pusuu | pusu |
| p Deletion | utuko | usuu | usu |
| Short *u → i/___# | -- | -- | usi |
| Vowel Harmony | -- | -- | isi- |
| Final Vowel Deletion | utuk | usu | -- |
| Surface Form | utuk | usu | isi- |

(Vowel harmony is allowed to apply twice here; therefore, it is assumed that this rule is older than p deletion and u fronting and still existed in the grammar at the time these rules applied.)

This may be a plausible way to account for these forms. Vowel harmony and the fronting of final short u are reasonably well established facts about the history of PNP phonology. Appeal to these rules will also be made when discussing prothetic w in Section 4.5. But, the ordering presented above is based on too few forms to be convincing. Further, those forms containing *t above are irregular in another respect. *t should have deleted. And, there still remains in the data one instance of *p deleting before *i that cannot be accounted for.

| | | |
|---|---|---|
| *api | aay | *fire* |

If *p does in fact sporadically delete before i, then the previous discussion is suspect.

### 4.4.2. *ngp

*ngp > pw

| | | |
|---|---|---|
| *ka(Rdr)angpa | kapw | *new* |
| *ngpongi | pwoong | *night* |

Only these two cases occur in the data.

### 4.4.3. *mp

*mp > pw

| | | |
|---|---|---|
| *mpo- | pwoo | *smell* |
| *mpou | mwpwey | *gall bladder* |
| *mpua | pwuu | *betel nut* |
| *mpule | pwili | *cowry shell* |
| *mputo | pwuus | *navel* |
| *mpengka | pwɛɛk | *bat* |
| *mpou | pwow | *fishing pole* |

Two exceptions occur:

| | | |
|---|---|---|
| *mponot | pon | *blocked* |
| *sampo | tapatap | *catch* |

No obvious explanation suggests itself for these irregular reflexes. They illustrate again the problem of oral/nasal grade contrast in POC.

### 4.4.4. *m

*m > m

| | | |
|---|---|---|
| *lima | lima- | *five* |
| *mada | maar | *preserved breadfruit* |
| *manu(k) | maan | *bird, animal* |
| *mata | maas | *eye* |
| *mauRi | mɛyng | *left* |
| *nsama | taam | *outrigger* |

| | | |
|---|---|---|
| *mansu(rR) | mɛt | *full* |
| *tama | saam | *father* |
| *ma?udi(p) | mowr | *life* |

*m > mw

| | | |
|---|---|---|
| *kamu | komw | *you* - honorific |
| *limu | liimw | *moss, sponge* |
| *-mu | -mw | *your* - sg |
| *?umu | uumw | *earth oven* |
| *namo | naamw | *lagoon* |

Thus, *m is reflected as mw before round vowels and as m elsewhere.

### 4.4.5. *ngm

| | | |
|---|---|---|
| *ndangma | ṭaamw | *forehead* |
| *ngmao | mwaang | *taro* |
| *ngmata | mwaas | *worm* |
| *Rungma(?) | iimw | *house* |

No exceptions occur here: *ngm > mw.

### 4.4.6. *w

*w > w

| | | |
|---|---|---|
| *awang | aaw | *mouth* |
| *nsiwa | tuwa- | *nine* |
| *walu | wɛli- | *eight* |
| *wawa | waawa | *newphew relationship* |
| *masawa | mataw | *open sea* |

In a single instance, *w is reflected as Ø.

| | | |
|---|---|---|
| *pakiwak | pako | *shark* |

This cognate will be further discussed in Section 5.1.1.

## 4.5. Glide insertion

Glides in PNP originate from three possible sources — from proto-glides (in the case of *w), from *i or *u, or from glide insertion. In this section of the paper, only those glides which occur as a consequence of insertion are to be considered.

As has already been suggested in Section 2.3., one glide insertion rule is the following.

$$\begin{bmatrix}+\text{syllabic}\\-\text{high}\end{bmatrix}\begin{bmatrix}+\text{syllabic}\\+\text{high}\\\alpha\text{back}\\\beta\text{round}\end{bmatrix}\#\Rightarrow\begin{bmatrix}+\text{syllabic}\\-\text{high}\end{bmatrix}\begin{bmatrix}-\text{syllabic}\\+\text{high}\\\alpha\text{back}\\\beta\text{round}\end{bmatrix}\begin{bmatrix}+\text{syllabic}\\+\text{high}\\\alpha\text{back}\\\beta\text{round}\end{bmatrix}\#$$

This rule inserts $\begin{bmatrix}y\\w\end{bmatrix}$ between a non-high vowel and $\begin{bmatrix}i\\u\end{bmatrix}$ #.

Examples follow:

| | | |
|---|---|---|
| *api | aay | *fire* |
| *mpou | pwow | *fishing pole* |
| *topu | seew | *sugarcane* |
| *sa?it | teytey | *sew* |
| *tanum | sowsow | *grave* |

That is, after the deletion of appropriate consonants, glides are inserted and the final vowel is deleted.

The principal basis of this argument is that within a generative framework it is synchronically necessary to preserve the final vowel of at least some of these forms in order to account for morphophonemic alternation. Thus aay *fire* occurs in a compound as kisi-ni-ɛy *a fire, as for cooking* (lit. bit-of-fire). Only syllabic i can raise short a to ɛ; therefore, a base ai must be postulated to account for the elicitation form, which undergoes compensatory lengthening, and the compound form which does not. Glide insertion is thus required synchronically; the extent to which it is necessary diachronically is unclear since, as was explained previously, it is not known to what extent final vowels have been preserved in the base.

In three instances the preceding rule seems to give the wrong result.

| | | |
|---|---|---|
| *mpou | mwpwey | *gall bladder (its)* |
| *tapu- | sɛwi | *conch shell trumpet* |
| *ponu | weey | *turtle* |

Note, however, that the first two forms are morphologically complex. mwpwey is possessed and sɛwi occurs in POC as a bound stem. It is not immediately obvious, though, how these affixes interact with the stem to produce the glide they do. The third form, *ponu > weey is aberrant in a number of respects, and remains without explanation.

Prothetic w also occurs as a consequence of glide insertion. Note these examples:

| | | |
|---|---|---|
| *eno | wen- | *lie down* |
| *onom | wene- | *six* |
| *ponu | weey | *turtle* |
| *?una | winɛ | *feathers, scales (its)* |
| *puki | wik | *turn* |

If one assumes that in the first form, *e was backed and rounded to o in harmony with the final vowel, that in the last three forms initial *p and *? were deleted (recall that *p deletes before round vowels), then one may postulate that w was inserted between # and round vowels. Later developments then unrounded these vowels.

Synchronically prothetic w is preserved only before non-round segments. Thus:

| | | |
|---|---|---|
| *puko | uuk | *net* |
| *pinsiko | utuk | *flesh* |
| *Ruap | uu | *tide* |
| *tupa | uup | *fish poison* |
| *pitu?u | usu | *star* |
| *?ulung | ulung | *pillow* |
| *punti | uut | *banana* |

There is still some instability in PNP between initial wi and initial u. The word for *lobster*, for example, is alternately pronounced wirɛnna or urɛnna, apparently in free variation.

## 4.6. Initial consonant clusters

There are four instances of initial consonant clusters in the data.

| | | |
|---|---|---|
| *muta(?) | mwmwus | *vomit* |
| *daRa(?) | nta | *blood* |
| *mpou | mwpwey | *gall bladder (its)* |
| *Ripa | mpɛ | *close (to him)* |

As Goodenough suggested in his article "The Long or Double Consonants of Trukese"[13] these occurrences of initial consonant clusters probably reflect older forms with first syllable reduplication. Comparing the first two items with Kiribati (Gilbertese) (KIR) and TRK, one finds:

| KIR | TRK | PNP |
|---|---|---|
| MuMuta | bbus | mwmwus |
| raraa | cca | nta |

TRK and PNP both delete the first vowel in the reduplicated syllable, thus producing geminate consonants. In PNP, however, still another rule applies. Synchronically, this rule is very general; in effect, if two adjacent homorganic consonants occur, the second of which is voiceless, the first becomes a homorganic nasal. This rule then accounts for nta from tta. What the precise nature of this rule was diachronically is not yet known.

## 4.7. A summary of regular consonantal reflexes

| | | | |
|---|---|---|---|
| Final Consonant Deletion | C | > | Ø/____# |
| Laryngeals | *? | > | Ø |
| Velars | {*k, *ngk} | > | k |
| | *ng | > | ng |
| | *R | > | Ø in seven cases<br>r in two cases |

| | | | |
|---|---|---|---|
| Coronals | {*s, *ns} | > | t |
| | {*d, *r} | > | r |
| | {*nd, *nt} | > | ṯ |
| | *t | > | Ø/___high vowels<br>s/elsewhere |
| | *n | > | n |
| | *l | > | l |
| Labials | p | > | Ø/___ round vowels<br>p/elsewhere |
| | {*ngp, *mp} | > | pw |
| | *m | > | mw/___round vowels<br>m/elsewhere |
| | *ngm | > | mw |
| | *w | > | w |
| Glide Insertion Rules | See Section 4.5. | | |

## 5. POC SOURCES FOR PNP VOWELS

Five vowels are postulated for POC. These are *i, *e, *a, *o, and *u. PNP reflexes of these POC vowels will be presented in the following discussion. First, the high vowels will be examined, then the mid vowels, and finally the low vowel a. As in the discussion of consonants, lists of regular correspondences will not be exhaustive, but lists of exceptions will.

### 5.1. High vowels

#### 5.1.1. *i

*i > i

| | | |
|---|---|---|
| *ia | ii | *he, she, it* - emphatic |
| *iriR | irip | *fan* |
| *ngkinit | kinii | *pinch* |
| *kinta | kiṯ- | *we* |
| *lima | lima- | *five* |
| *pinta | piiṯ | *split leaves of pandanus* |
| *piti(k) | piṯik | *spring up* |
| *tina | iin | *mother* |
| *(dl)iki | ṯik | *small* |
| *anitu | ɛni | *ghost* |

| | | |
|---|---|---|
| *iku | iik | *tail* |
| *limu | liimw | *moss, sponge* |
| *niu | nii | *coconut* |
| *pitu | isi- | *seven* |

*i > ɛ

| | | |
|---|---|---|
| *talinga | salɛng | *ear* |

In this instance, *i → ɛ/aC___Ca. This rule can also be used to explain the following odd form.

| | | |
|---|---|---|
| *pakiwak | pako | *shark* |

After the application of the preceding rule and the deletion of the final consonant and vowel, the intermediate form pakɛw occurs. A synchronic fact of PNP is that ɛw alternates with ow and sometimes o. This alternation, then, might account for the form pako.

In two instances, *i > u.

| | | |
|---|---|---|
| *pinsiko | utuk | *flesh* |
| *pitu?u | usu | *star* |

A possible explanation for why u occurs in these forms, but not in a form like isi- from *pitu, has already been presented in Section 4.4.1. under the discussion of the loss of *p.

In a single instance, *i > Ø.

| | | |
|---|---|---|
| *inu(m) | nim | *drink* |

### 5.1.2. *u

*u > u

| | | |
|---|---|---|
| *tu?u(d) | uu | *stand* |
| *?ulunga | ulung | *pillow* |
| *muta(?) | mwmwus | *vomit* |
| *tupa | uup | *fish poison* |
| *mpua | pwuu | *betel-nut* |
| *?uda | urɛ(nna) | *lobster* |
| *puko | uuk | *net* |
| *mputo | pwuus | *navel* |
| *uto | uus | *float of fishing net* |
| *nuns(io) | nuut | *squid* |

Note that in these examples, *u > u in those cases where the vowel is inherently long, as in the first form after the deletion of *?, or where the stem-final vowel is *a or *o. This observation lends further credibility to the argument developed in Section 4.4.1., and referred to above, where it was hypothesised that high back vowels remained back in stems ending in short non-high vowels.

In cases where the final vowel of the reconstructed form is high, the situation is more complex. Note these examples where the final vowel is *i.

| | | |
|---|---|---|
| *kuli | kiil | *skin* |
| *puki | wik | *turn* |
| *suli | iil | *shoot, sucker* |
| *?uling | iliil | *steer* |
| *mudi | mwiri | *after (it)* |

Here, *u is reflected as i. This can be attributed to vowel harmony of the type previously discussed in this paper.

Now note the following forms where the final vowel is *u.

| | | |
|---|---|---|
| *kuku | kiik | *nail* |
| *matudu | mɛyr | *sleep* |
| *mpulu(t) | pwiil | *gum, glue* |
| *sulu | tiil | *torch* |
| *tumpu(?) | ipw | *parentage* |
| *tusu(k) | it | *point* |

If one assumes that short *u → i/___#, then vowel harmony will again explain the fronting of non-final u to i. Based on synchronic evidence, the fronting of final *u is a reasonable hypothesis. (See Section 3.2.). Further support for the position that short *u was fronted to i in final position is provided by forms like these.

| | | |
|---|---|---|
| *walu | wɛli- | *eight* |
| *pːtu | isi- | *seven* |

In these forms, which were free in POC but only occur affixed in PNP, the final vowel is clearly preserved as i.

In two cognates, *u fronts to i even though it is not in final position.

| | | |
|---|---|---|
| *inu(m) | nim | *drink* |
| *pang(ou)(n) | pangin | *awaken* |

Two explanations suggest themselves here. It may be that *u fronts to i even if an intervening non-round segment occurs, or more likely, the final consonants here, which should have deleted, are analysed as being thematic and not part of the stem itself.

All the preceding high vowel reflexes, then, might be accounted for by these two rules.

Vowel Harmony

$$\begin{bmatrix} +\text{syllabic} \\ -\text{low} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\text{back} \\ \beta\text{round} \end{bmatrix} / \underline{\quad} \; [+\text{cons}] \begin{bmatrix} +\text{syllabic} \\ -\text{low} \\ \alpha\text{back} \\ \beta\text{round} \end{bmatrix}$$

*u Fronting

$$\begin{bmatrix} +\text{syllabic} \\ +\text{high} \\ +\text{back} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{back} \\ -\text{round} \end{bmatrix} / \underline{\quad}\#$$

Vowel harmony is probably the older of these two rules, and very likely applied both before and after the fronting of *u. Again, I refer the reader to Section 4.4.1. for examples of this position.

As one would expect, there are exceptions to these rules.

| | | |
|---|---|---|
| *ʔumu | uumw | *earth oven* |
| *muntu | mwut' | *pulverised* |
| *punti | uut' | *banana* |
| *tuki | suk | *pound* |
| *Rungma | iimw | *house* |
| *dua | ria- | *two* |
| *nsiwa | tuwa- | *nine* |

In the first four forms (assuming the fronting of final *u) the vowel harmony rule does not apply. Even synchronically, though, vowel harmony (or copying) is a very complex phenomenon, the precise nature of which is not understood. The fronting of *u in the fifth and sixth forms is also unexplainable. The last form, though, might be explained by the fact that iw is synchronically an impermissable sequence; thus, *i was probably backed and rounded to u as a consequence of this constraint.

### 5.1.3. Glide formation

POC high vowels become glides in the following items.

| | | |
|---|---|---|
| *matudu | mɛyr | *sleep* |
| *maʔudi(p) | mowr | *life* |
| *ngapulu | ngowl | *ten* |
| *pua(ʔ) | waa | *fruit* |

A comparison of the first three forms suggests that rules affecting *t are more recent than the loss of *ʔ or *p. Assuming a glide formation rule

$$\begin{bmatrix} +\text{syllabic} \\ +\text{high} \end{bmatrix} \rightarrow [-\text{syllabic}] / \begin{bmatrix} +\text{syllabic} \\ -\text{high} \end{bmatrix} \_\_\_\_ [-\text{syllabic}]$$

then the following rule ordering would explain these forms.

| POC Reconstructions | matudu | maʔudi | ngapulu |
|---|---|---|---|
| *ʔ → ø | -- | maudi | -- |
| Glide Formation | -- | mawdi | -- |
| *p → ø/___round V | -- | -- | ngaulu |
| Glide Formation | -- | -- | ngawlu |
| *u → i/___# | matudi | -- | ngawli |
| Vowel Harmony | matidi | -- | -- |
| *t → ø/___high V | maidi | -- | -- |
| Glide Formation | maydi | -- | -- |
| Other Rules | mɛyr | mowr | ngowl |

Note that the glide formation rule has been allowed to apply whenever the conditions are correct. This, in essence, is claiming that this is a rule which has been in the language for a long time. Also note that the ordering here is consistent with ordering arguments presented elsewhere in this paper. It should not be assumed, however, that rules subsumed under 'Other Rules' are ordered in this position. Rather, since these rules are not directly relevant to the argument, they are grouped together to avoid obscuring the presentation.

The fourth form, waa from *pua(ʔ), can be explained by a glide formation rule of this nature.

$$\begin{bmatrix}+\text{syllabic}\\+\text{high}\end{bmatrix} \rightarrow [-\text{syllabic}] \,/\, \#\_\_\_ \begin{bmatrix}+\text{syllabic}\\-\text{high}\end{bmatrix}$$

If one takes the position that juncture has a minus value for all distinctive features that apply to segments, then the two glide formation rules presented here can be written as a single mirror image rule.

$$\begin{bmatrix}+\text{syllabic}\\+\text{high}\end{bmatrix} \rightarrow [-\text{syllabic}] \quad \% \quad \begin{bmatrix}+\text{syllabic}\\-\text{high}\end{bmatrix} \_\_\_\_ [-\text{syllabic}]$$

## 5.2. Mid vowels

### 5.2.1. *e

There are relatively few non-final occurrences of *e.

| | | |
|---|---|---|
| *eno | wen(ti) | *lie down* |
| *kiekie | kiɛ | possessive for *mat* |
| *mpengka | pwɛɛk | *bat* |

These forms support the position taken in Section 2.3. that *e is realised as e when followed by a high non-consonantal segment, and that elsewhere *e > ɛ. The first item, for example, when followed by the suffix -to, is pronounced wɛnto.

In a single instance, *e > a.

| | | |
|---|---|---|
| *katea | kasa | *side of canoe opposite outrigger* |

Based on a single example, one might hypothesise that ea > aa.

There are additionally two questionable cognates reflecting *e.

| | | |
|---|---|---|
| meRa | mer | *rust* |
| pengu | pɛngi(tɔy) | |
| | pangit | *blow the nose* |

The construct form of the first word is merin, thus reflecting the wrong final vowel (and explaining the occurrence of e rather than ɛ.) In the second item, a better reconstruction for the PNP form would seem to be *pangu.

### 5.2.2. *o

*o > o

| | | |
|---|---|---|
| *(dr)on(st)o | rot | *dark* |
| *dongo | rong | *hear* |
| *mpo- | pwoo | *smell* |
| *ngpongi | pwoong | *night* |
| *mpou | pwow | *fishing pole* |
| *nsoka | tok | *stab* |
| *toko(n) | sokon | *cane* |

*o > e/ɛ

| | | |
|---|---|---|
| *topu | seew | *sugarcane* |
| *ponu | weey | *turtle* |
| *ko(e) | kɛ | *you* - sg |
| *mpou | mwpwey | *gall bladder (its)* |
| *onom | wene- | *six* |

The fronting of *o to e or ɛ apparently involves vowel harmony, but since the precise nature of vowel harmony is not understood, it is difficult to account for these forms systematically. The last cognate also suggests that final *o was fronted to e, but this cannot be justified by synchronic data. More likely, final *o was lost; for example, the vowel that appears in the construct form of a noun like roong *news* (from *dongo *hear*) is just an insert vowel. Thus, one may say either rongin or rongɛn. Both forms are acceptable.

In two instances, *o becomes a high vowel.

| | | |
|---|---|---|
| *tolu | sili- | *three* |
| *matolu | mosul | *thick* |

In both instances, *o is between s and l, which are phonetically palatalised. This palatalisation probably accounts for the raising, but the motivation for the front/back differences here is again not very clear. In general, reflexes of mid vowels are not very well understood, principally because there are relatively few forms to work with.

## 5.3. *a

*a > ɔ

| | | |
|---|---|---|
| *angoango | ɔngɔɔng | *yellow* |
| *ʔatop | ɔɔs | *thatch* |
| *lango | lɔɔng | *fly* |

These three forms illustrate what was apparently an early development affecting *a; that is, *a → ɔ/___Co. Note that this rule precedes compensatory lengthening. (The long vowel in ɔngɔɔng is a consequence of reduplication rather than lengthening.) Unfortunately, there are only three examples of this rule, and almost as many exceptions. Thus:

| | | |
|---|---|---|
| *namo | naamw | *lagoon* |
| *sampo | tapatap | *catch* |

The second of these, though, may not be a good cognate; *mp should have been reflected as a velarised, not a plain, labial.

In the following forms, *a is reflected as a.

| | | |
|---|---|---|
| *-da | -ra- | *their* |
| *daRa(?) | nt̪a | *blood* |
| *dua | ria- | *two* |
| *kaka | kaaka | *respect* |
| *lima | lima- | *five* |
| *madama | maram | *moon* |
| *manang | manaman | *spiritual power* |
| *masawa | mataw | *sea* |
| *pua(?) | waa | *fruit* |
| *-nta | -t̪a- | *our* |

Note that in all of the above PNP cognates a is short, except as a consequence of reduplication or compensatory lengthening. Note also that in the POC reconstructions, *a is followed by no vowel except *a.

In these forms, after the deletion of a final consonant, the final vowel is *i. Note that *a raises to ɛ.

| | | |
|---|---|---|
| *manipi(s) | mɛnipinip | *thin* |
| *masaki(t) | mɛtɛk | *pain* |
| *pali | pɛl | *respect* |
| *mauRi | mɛyng | *left hand* |
| *sa?it | tɛɛk | *sew* |
| *tangi(s) | sɛng | *weep* |

The following forms are all possessed nouns. Synchronically the third person possessive suffix is -i.

| | | |
|---|---|---|
| *?apara | apɛrɛ | *shoulder (its)* |
| *asang | ɛtɛ | *gill (its)* |
| *pan(i) | pɛɛ | *wing (its)* |

Thus, the raising here might be attributed to the presence of an -i suffix, which is deleted on the surface.

The following items also exhibit the raising of *a to ɛ, but the final vowel is *u.

| | | |
|---|---|---|
| *anitu | ɛni | *ghost* |
| *mansu(rR) | ɪmɛt | *full* |
| *matudu | mɛyr | *sleep* |
| *natu | nɛɛ | possessive for *child* |
| *naskaRu | tɛkɛ | *island* |
| *walu | wɛli- | *eight* |

If the raising of a to ɛ is ordered after the fronting of u to i, then these forms, too, illustrate regular reflexes of *a. This a raising rule can now be formulated as follows:

$$\begin{bmatrix} +\text{syllabic} \\ +\text{low} \\ -\text{long} \\ -\text{back} \end{bmatrix} \rightarrow [-\text{low}] \,/\, \_\_ C_0 \begin{bmatrix} +\text{syllabic} \\ +\text{high} \\ -\text{back} \end{bmatrix} \#$$

This rule states that short a (probably a front vowel at this point as a consequence of the introduction of ɔ) raises to ɛ when followed by any number of consonants and i in final position. Final position appears to be important based on forms like this one.

| | | | |
|---|---|---|---|
| *pang(ou)(n) | pangin | trans. | |
| | mpɛng | intr. | *awaken* |

There are diachronically few cases which support the importance of final position, but synchronically it is clearly important. Additionally, a vowel copying rule is needed to account for the raising of the first a in forms like ɛtɛ *gill (its)*, mɛtɛk *pain*, etc. (See Section 2.3. for further comments on this rule.)

The following items illustrate that a raising should also be ordered after compensatory lengthening.

| | | |
|---|---|---|
| *api | aay | *fire* |
| *ansa(n) | aat | *name* |
| *awang | aaw | *mouth* |
| *da?a(n) | raa | *branch, twig* |
| *laki | laak | *penis of adult* |
| *langi(t) | laang | *sky* |
| *manu(k) | maan | *bird, animal* |
| *namu | aamw | *bug* |
| *pani | paan | *bait* |
| *sala(n) | aal | *road* |
| *nsama | taam | *outrigger* |
| *tali | saal | *rope, cord* |
| *tama | saam | *father* |

Note that a raising affects only short vowels. Thus, if this rule is ordered after lengthening, even if a final high vowel occurs, the rule cannot apply.

There are instances, though, where a raises to o under the influence of final *u.

| | | |
|---|---|---|
| *kamu | komw | *you* - honorific |
| *ngapulu(?) | ngowl | *ten* |
| *tanum | sowsow | *grave* |
| *tapu- | sowi | *conch shell trumpet* |
| *tau | sow- | *practitioner of* |
| *ta?u | sow(npar) | *year* |

In each of these examples, note that o precedes either a velarised consonant or w. It is difficult to know just what the significance of this is. It may be that *u did not front in this environment, and thus raised *a to o, or possibly, where *u was final it did front with subsequent raising of *a to ɛ, but ɛ backed to o before mw or w.

In two cases, *a is reflected as e.

| | | |
|---|---|---|
| *tansi(k) | seet | *sea* |
| *pa- | pey | *fight* |

Apparently this vowel occurs as a consequence of restructuring of the lexicon. The most common word for *sea* in PNP is nanset (lit. in-sea). The elicitation

form seet was probably formed from the compound form; thus set after compensatory lengthening became seet. pey has no related forms where alternation would provide evidence for *a. Thus, synchronically perhaps these two forms occur in the base as sɛti and pɛi. Diachronically, then, the raising of a to ɛ occurred, and because synchronically ɛ is the vowel of the base and a high segment follows, ɛ raises to e.

These unexplained cognates remain:

| | | |
|---|---|---|
| *mate | mɛɛ(la) | *die* |
| *mapo | mo | *heal* |
| *sau(?) | too | *far off* |
| *?ate | ɛɛ | *liver* |
| *ndau(n) | tɛɛ | *leaf* |
| *mataku(t) | masak | *afraid* |
| *matolu | mosul | *thick* |
| *ka(Rdr)angpa | kapw | *new* |

In each an unexpected reflex of *a occurs, either in terms of height, frontness, or length. Possible explanations could be offered, but they would be *ad hoc* to an extreme.

## 5.4. A summary of regular vowel reflexes

Only regular reflexes of POC vowels will be listed here. Conditioning will be stated in terms of the rules that have been discussed.

High Vowels

| | | | |
|---|---|---|---|
| *i | > | u | Vowel Harmony |
| | | ɛ | i Lowering |
| | | y | Glide Formation |
| | | i | Elsewhere |
| *u | > | i | u Fronting and Vowel Harmony |
| | | w | Glide Formation |
| | | u | Elsewhere |

Mid Vowels

| | | | |
|---|---|---|---|
| *e | > | e | ɛ Raising |
| | | ɛ | Elsewhere |
| *o | | e/ɛ | Vowel Harmony (?) |
| | | o | Elsewhere |

The Low Vowel *a

| | | | |
|---|---|---|---|
| *a | > | ɔ | a Backing |
| | | ɛ | a Raising |
| | | o | Conditioning not clear |
| | | a | Elsewhere |

Other Rules Affecting Vowels

Compensatory Lengthening
Final Vowel Deletion

## 6. CONCLUDING COMMENTS

The purpose of this paper has been to present some preliminary observations on the history of PNP phonology. Like most studies of this nature, however, the comments have raised perhaps as many questions as they have answered. Certainly, there are many forms in PNP that appear to be cognate with POC that cannot be systematically accounted for by the sound change rules that have been proposed. These difficulties cannot be eliminated, but their possible sources ought to be considered, for they suggest future areas of research.

Obviously, a better understanding of synchronic phonological phenomena in PNP would lead to improved diachronic hypotheses. The restraints on vowel harmony, for example, and the extent of final vowel retention, if known, would almost certainly eliminate some of the difficulties involved in establishing PNP reflexes for POC vowels.

Arguments concerning rule ordering, too, need to be better developed and tested. In this paper, a number of arguments concerning ordering have been suggested, but unfortunately, these too often are based on only a few cognates. Ideally, a list of POC reconstructions should be presented in conjunction with a sequence of ordered rules which would correctly predict current PNP forms.

Probably another source of difficulty is that false cognates have been taken into consideration in formulating sound change rules. Because of the preliminary nature of this study, however, all items which might reasonably be considered cognate have been included. The elimination of some of the more troublesome forms would certainly produce neater results, but it is impossible at this stage of research, and perhaps at any stage, to establish precise criteria that would be necessary to insure that only true cognates were considered.

## NOTES

1. I am indebted to Mr. Damian Sohl who provided the date for this chapter. I also wish to thank him, along with George Grace, Irwin Howard, and fellow participants in the Micronesian Seminar, for the many helpful comments I received concerning the analysis of these data. I, of course, accept final responsibility for all errors.

2. Other languages spoken in Ponape District are Pingelapese, Mokilese (MOK), Kosraean (Kusaiean) (KSR), Nukuoro, and Kapingamarangi. The first three of these are Micronesian; the last two are Polynesian. For further discussion of the relations of PNP to other MC languages, see Rehg 1981:3-12.

3. The adoption of the Main dialect as a basis for spelling was made official in January, 1972 at an orthography conference conducted by the author. In practice, this procedure is much older.

4. Garvin, n.d., p.8-9.

5. A taxonomic inventory of PNP consonants would be identical.

6. Garvin, n.d., p.5.

7. A single l-r exception occurs; that is luhr *shrimp*. Four n-ng exceptions have been found, but these are all of the canonical form ..ngVn.. . The significance of this is not clear.

8. The construct form ni occurs without the final vowel being deleted in compounds like kisi-ni-ɛy (lit. bit-of-fire) meaning *a fire, as for cooking.*

9. Grace, (Micronesian Seminar, Fall 1971, Spring 1972).

10. Micronesian Seminar, Spring 1972.

11. Gulick, 1880:97.

12. Damian Sohl, a native speaker of PNP, reports having heard older speakers of PNP still using the voiceless palatal affricate. (personal communication)

13. Goodenough, 1963:78.

## APPENDIX

### Proto-Oceanic/Ponapean Cognates

Reconstructions based on Dyen (1949) will be preceded by D; those from Blust (1972) will be preceded by B. All others are from Grace (1969).

| | POC | PNP |
|---|---|---|
| 1. | -ʔaki cause or instrument suffix | -ki instrumental suffix |
| 2. | alap *take* | alɛ *take* |
| 3. D | anitu *ghost* | ɛni *ghost* |
| 4. | ʔapaRa *shoulder* | apɛrɛ *shoulder (its)* |
| 5. | api *fire* | aay *fire* |
| 6. | ansa(n) *name* | aat *name* |
| 7. | asang *gill, fin* | ɛtɛ *gill (its)* |
| 8. | ʔate *liver* | ɛɛ *liver (its)* |
| 9. | ʔatop *thatch, sago palm* | ɔɔs *thatch, sago palm* |
| 10. | awang *open space, mouth* | aaw *mouth* |
| 11. | -da *their* | -ra- *their* |
| 12. | daʔa(n) *branch, twig* | raa *branch, twig* |
| 13. | ndangma *forehead* | ṯaamw *forehead* |
| 14. | ndamu *chew, lick* | ṯamwɛ *lick* |
| 15. | (dr)an(i) *day* | raan *day* |
| 16. | daRa(ʔ) *day* | nṯa *blood* |
| 17. | ndau(n) *leaf* | ṯɛɛ *leaf* |
| 18. | (dl)iki *small* | ṯik *small* |
| 19. | (dr)on(st)o *night* | roṯ *dark* |
| 20. | dongo *hear* | rong *hear* |
| 21. | dua *two* | ria- *two* |
| 22. | eno *lie down* | wen(ti) *lie down* |
| 23. | iku *tail* | iik *tail* |
| 24. | inu(m) *drink* | nim *drink* |
| 25. | iriR *to fan* | irip trans.<br>irir intr. *to fan* |
| 26. | ka(dr)u *scratch* | karu(pwut) *scratch* |
| 27. | ka(Rdr)angpa *new* | kapw *new* |
| 28. | kaka *older sibling of same sex* | kaaka *respect of elders* |
| 29. | kamu *ye* | komw *ye* - honorific |
| 30. | katea *side of canoe opposite outrigger* | kasa *side of canoe opposite outrigger* |
| 31. | kiekie *pandanus* | kiɛ poss. *for mats* |
| 32. | ngkinit *pinch* | kinii trans.<br>kinikin intr. *pinch* |
| 33. | kinta 1st incl | kiṯ- 1st pl stem |

| | | POC | PNP |
|---|---|---|---|
| 34. | | ko(e) *thou* | kɛ *you - sg* |
| 35. | | kuku *claw, nail* | kiik *nail* |
| 36. | | kuli(t) *skin, bark* | kiil *skin, bark* |
| 37. | | kuRita *octupus* | kiis *octupus* |
| 38. | | laki *male, adult* | laak *penis of adult* |
| 39. | | langi(t) *sky* | laang *sky* |
| 40. | | lango *a fly* | lɔɔng *a fly* |
| 41. | B | (ln)ima *bail* | lim *bail* |
| 42. | | lima *five* | lima- *five* |
| 43. | | limu *seaweed, moss* | liimw *moss, sponge* |
| 44. | | maa *be ashamed* | maak *be ashamed* |
| 45. | | mada *fermented, ripe* | maar *preserved breadfruit* |
| 46. | D | madama *moon* | maram *moon* |
| 47. | | ngmalala *cleared ground* | mall *clear grassy spot* |
| 48. | | manang *power, wind* | manaman *spiritual power* |
| 49. | | manipi(s) *thin* | mɛnipinip *thin* |
| 50. | | manu(k) *bird, animal* | maan *bird, animal* |
| 51. | | ngmao *taro* | mwaang *taro* |
| 52. | | mapo *to heal* | mo *to heal* |
| 53. | | masaki(t) *sick, pain* | mɛtɛk *pain* |
| 54. | | masawa *shore, sea* | mataw *open sea* |
| 55. | | mata *eye* | maas *eye* |
| 56. | | ngmata *snake, worm* | mwaas *worm* |
| 57. | | mataku(t) *afraid* | masak *afraid* |
| 58. | | mansu(rR) *plenty of food* | mɛt *full of food* |
| 59. | | mate *die* | mɛɛ(la) *die* |
| 60. | | matolu *thick* | mosul *thick* |
| 61. | | matudu(R) *sleep* | mɛyr *sleep* |
| 62. | | mauRi *left (hand)* | mɛyng *left* |
| 63. | | ma?udi(p) *life* | mowr *life* |
| 64. | | meRa *reddish* | mer *rust* |
| 65. | | -mu *your - sg* | -mw *your - sg* |
| 66. | | mudi *behind* | mwiri *behind, after (it)* |
| 67. | | muta(?) *to spit, vomit* | mwmwus *vomit* |
| 68. | | muntu *severed, cut short* | mwut *anything pulverised* |
| 69. | | namo *lagoon* | naamw *lagoon* |
| 70. | | namu(k) *mosquito* | aamw *bug* |
| 71. | | nana(?) *pus* | naan *pus* |
| 72. | | niu(R) *coconut* | nii *coconut* |
| 73. | | natu *child, offspring* | nɛɛ poss *for child* |
| 74. | | nuns(io) *squid* | nuut *squid* |
| 75. | | ngapulu(?) *ten* | ngowl *ten* |
| 76. | D | ngipo *tooth* | ngii *tooth* |
| 77. | | onom *six* | wene- *six* |
| 78. | | pa- *fight* | pey *fight* |
| 79. | | pa- causative prefix | pa- causative prefix |
| 80. | | pa?a *stem of ensiform leaves* | pa *leaf including stem* |
| 81. | | pa?a *cut yams for planting* | pa(ki) *cut yams for planting* |
| 82. | | pada *pandanus* | (ki)par *pandanus* |
| 83. | | pakiwak *shark* | pako *shark* |
| 84. | D | pali *respect* | pɛl *respect* |
| 85. | | pan(i) *wing* | pɛɛ *wing, arm (its)* |
| 86. | | pani *bait* | paan *bait* |

| | POC | PNP |
|---|---|---|
| 87. | papa *downwards* | paa *downwards* |
| 88. | pang(ou)(n) *awaken* | pangin *awaken* |
| 89. | pat *four* | paa- *four* |
| 90. | mpengka *bat* | pwɛɛk *bat* |
| 91. | pengu *blow the nose* | pangit trans.<br>pɛngi(ʈɔy) intr. *blow the nose* |
| 92. | pi(dr)i *fold, twist* | pir *spin, turn* |
| 93. | pili(?) *select, choose* | pil *choose* |
| 94. | pinsiko *flesh* | utuk *flesh* |
| 95. | pinta *split* | piiʈ *split leaves of pandanus* |
| 96. | piti(k) *spring up* | piʈik trans.<br>piʈ intr. *spring up* |
| 97. | pitu *seven* | isi- *seven* |
| 98. | pitu?u *star* | usu *star* |
| 99. | mpo- *to smell* | pwoo *smell* |
| 100. B | mponot *blocked* | pon *blocked* |
| 101. | ponu *turtle* | weey *turtle* |
| 102. | ngpongi *night* | pwoong *night* |
| 103. | pua(?) *fruit* | waa *fruit* |
| 104. | mpua *areca nut* | pwuu *areca nut* |
| 105. | mpule *cowry shell* | pwili *cowry shell* |
| 106. | puki *return* | wikit trans.<br>wik intr. *to turn* |
| 107. | mpulu(t) *gum, glue* | pwiil *gum, glue* |
| 108. D | puko *net* | uuk *net* |
| 109. D | punti *banana* | uuʈ *banana* |
| 110. | mputo *navel* | pwuus *navel* |
| 111. | raku *scratch* | raki trans.<br>rɛkirɛk intr. *scratch* |
| 112. | Ripa *go close* | mpɛ *close (to him)* |
| 113. | Ruap *rising tide* | uu *tide* |
| 114. | Rungma(?) *house* | iimw *house* |
| 115. | sa?it *sew* | tɛɛk trans.<br>teytey intr. *sew* |
| 116. | sakan *cooking and everything associated with it* | sakan trans.<br>sak intr. *eat* - honorific |
| 117. | sala(n) *road, path* | aal *road, path* |
| 118. | nsama *outrigger* | taam *outrigger* |
| 119. | sampo *catch* | tapɛr trans.<br>tapatap intr. *catch* |
| 120. | sanga *thigh* | taang *thigh* |
| 121. | sangka(?) *to hop* | takɛ *ride (a vehicle)* |
| 122. | nsangi *wind* | aang *wind* |
| 123. | nsakaRu *reef, sandbank* | tɛkɛ *island* |
| 124. | nsake *upwards* | ta *upwards* |
| 125. | nsapa *what* | ta *what* |
| 126. | sau(?) *outside, far off* | too *far off* |
| 127. | (s)ida *they* | ir- *they* |
| 128. | nsipo *downwards* | -ti *downwards* |
| 129. | nsiwa *nine* | tuwa- *nine* |
| 130. | nsoka *stab* | tok *stab* |

| | POC | PNP |
|---|---|---|
| 131. | solo *mountain* | tool *mountain* |
| 132. | sulu *torch* | tiil *torch* |
| 133. | sili *penetrate* | til *penetrate* |
| 134. B | suli *shoot, sucker* | iil *shoot, sucker* |
| 135. | susu *breast* | tiiti *breast* |
| 136. | -nta *our* (incl) | -ta- lst pl poss |
| 137. | tali *cord, rope* | saal *cord, rope* |
| 138. | talinga *ear* | salɛng *ear* |
| 139. | tama *father* | saam *father* |
| 140. | tanum *bury* | sowsow *grave* |
| 141. | tangi(s) *weep* | sɛng *weep* |
| 142. | tapu- *conch* | sɛwi or sowi *conch shell trumpet* |
| 143. B | tapu(n)i *afterbirth* | sow *afterbirth* |
| 144. | tansi(k) *sea, salt water* | seet *sea* |
| 145. | tau *man, person* | sow- *practitioner of* |
| 146. | ta?u *year, season* | sow(npar) *year* |
| 147. | tina *mother* | iin *mother* |
| 148. | toko(n) *staff, pole* | sokon *cane* |
| 149. | tolu *three* | sili- *three* |
| 150. | topu *sugarcane* | seew *sugarcane* |
| 151. | tuki *hammer, pound* | suk *hammer, pound* |
| 152. | tupa *fish poison* | uup *fish poison* |
| 153. | tumpu(?) *grow* | ipw- *parentage* |
| 154. | tusu(k) *to point* | it *to point* |
| 155. | tu?u(d) *stand up* | uu(ta) *stand up* |
| 156. | ?uda(ng) *lobster* | urɛ(nna) or wirɛ(nna) *lobster* |
| 157. | ?ulung *pillow* | ulung or uluul *pillow* |
| 158. | ?umu *earth oven* | uumw *earth oven* |
| 159. | ?una(p) *fish-scale* | winɛ *fish-scale (its)* |
| 160. | uto *float of fishing net* | uus *float of fishing net* |
| 161. | walu *eight* | wɛli- *eight* |
| 162. | wawa *mother's brother* | waawa *nephew relationship* |

## BIBLIOGRAPHY

BLUST, Robert A.

1972 Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list. *Working Papers in Linguistics, University of Hawaii* 4/1:1-43.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

1965 *A sketch of Trukese grammar.* Essay 4. New Haven: American Oriental Society.

GARVIN, Paul L.

n.d. (1951) A linguistic study of Ponape. MS. CIMA, 1947-49.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The proceedings of the Ninth Pacific Science Congress, 1957*, vol.3:77-86.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

n.d. Some Trukese cognates (from Dyen 1949). Mimeo. University of Hawaii.

GULICK, Luther H.

1880 A vocabulary of the Ponape dialect: Ponape-English and English-Ponape. *Journal of the American Oriental Society* 10:1-109.

REHG, Kenneth L.

1981 *Ponapean reference grammar*. Honolulu: University Press of Hawaii.

# NASAL SUBSTITUTION RULES IN PONAPEAN

Kenneth Rehg

## 1. INTRODUCTION[1]

In Ponapean (PNP)[2], a nuclear Micronesian (MC) language[3] spoken in the Eastern Caroline Islands, the two optimal consonant cluster types are those involving (1) geminate sonorants or (2) a sequence of a nasal followed by a homorganic obstruent. An inventory of the consonantal phonemes of PNP is presented in the chart below.[4]

Consonantal Phonemes

| | Labial | Dental | Alveolar | Post-Alveolar | Velar |
|---|---|---|---|---|---|
| Stops | | | | | |
| Plain | p | t | | | k |
| Velarised | $p^w$ | | | | |
| Fricatives | | | $s^y$ | | |
| Affricates | | | | ts | |
| Nasals | | | | | |
| Plain | m | n | | | ŋ |
| Velarised | $m^w$ | | | | |
| Liquids | | l | r̃ | | |

For orthographic convenience, these phonemes will subsequently be written as follows.

| The Phoneme | Will be Written |
|---|---|
| $p^w$ | pw |
| $m^w$ | mw |
| r̃ | r |
| ts | t' |
| ŋ | ng |

The remaining phonemes will be represented as they are on the preceding chart.

In accord with these orthographic practices, the optimal consonant clusters of PNP may be listed as follows.

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 317-337. *Pacific Linguistics*, C-80, 1984.

Optimal Consonant Clusters

| Geminate Sonorants | Nasal-Obstruent Sequences |
|---|---|
| mm | mp |
| mwmw | mwpw |
| nn | nt |
| ll | ns |
| rr | nt' |
| ngng | ngk |

The clusters ns and nt' are homorganic as a result of a process which assimilates n to the position of a following coronal obstruent.[5]

Consonant clusters not of these two types, non-optimal clusters, either do not occur or are subject to modification by (1) the insertion of an epenthetic vowel, (2) the complete assimilation of the first consonant to the second, or (3) the substitution of a nasal for the first of the two consonants in the cluster.[6] The focus of this paper is on the latter phenomenon, which will subsequently be referred to as nasal substitution.

## 2. SYNCHRONIC NASAL SUBSTITUTION RULES

The first description of nasal substitution in PNP was provided by Paul Garvin (1962:120). He observed:

> Morphemes with final p, t, k have within the same phrase sandhi variants dissimilated to final m, n, ŋ respectively before morphemes with initial consonant identical to their non-contact final....

Among the examples Garvin cited were the following, rewritten here in accord with the transcription system used in this paper.

| Morphemes | Pronounced |
|---|---|
| soop+piir | sompiir |
| *ship+fly* | *aeroplane* |
| mwoot+ti | mwoonti |
| *sit+down* | *sit down* |
| tɔtɔɔk+ki | tɔtɔɔngki |
| *work*+instrumental suffix | *work with* |

Garvin's account of nasal substitution in PNP, however, is deficient in three respects. First, data not considered by Garvin illustrate that other clusters in addition to geminate p, t, and k may undergo nasal substitution. Second, the condition that these clusters be in the same phrase is not sufficient to explain the constraints on the operation of nasal substitution. Third, nasal substitution is apparently best understood, not as a consequence of dissimilation, but rather of weakening.

The third point above will be explored in Section 4. of this paper, which deals with the motivation for nasal substitution. The first two points will be examined in the remainder of this section, where it will be argued that there are in fact two synchronic rules of nasal substitution, one of which is constrained in its application to reduplicated forms, while the other has a wider domain of application. The nasal substitution rule found in reduplicated forms will be examined first.

## 2.1. Nasal substitution in reduplication

Reduplication is productively employed in PNP to signal durative aspect.[7] There are at least eleven distinctive surface patterns of reduplication, the occurrence of which is governed by the phonological shape of the word being reduplicated. Two of these patterns, affecting words in which the first three segments are CVC, may lead to consonant clusters, as illustrated by the following examples:

Pattern I: Total Reduplication

| | |
|---|---|
| rer | *tremble* |
| rerrer | *trembling* |

Pattern II: Partial Reduplication

| | |
|---|---|
| rere | *peel* |
| rerrere | *peeling* |

Within the framework of generative phonology, both of these patterns may be characterised as initial CVCV reduplication. Evidence for this analysis follows.

As noted in Rehg 1973, there is a synchronic rule in PNP that deletes the final vowel of a polysyllabic base before word boundary; elsewhere, this vowel is retained. Thus, a base of the shape *$C_1V_1C_2V_2$ will surface as $C_1V_1C_2$ (unreduplicated) or as $C_1V_1C_2V_2C_1V_1C_2$ (reduplicated). Examples follow, where the final base vowel (or a conditioned variant of this vowel) is underlined in the reduplicated form.[8]

| Word | Gloss | Reduplicated |
|---|---|---|
| tep | *kick* | tepetep |
| tep | *start* | tepitep |
| kos | *throw* | kosokos |
| kos | *bent* | kosikos |
| net | *smell* | netenet |
| net' | *sell* | net'inet' |
| let' | *flick* | let'elet' |

The base form of an intransitive verb like tep *kick*, therefore, is *tepe (cf. the transitive form tepek). The surface form tepetep is derived as follows:

| | |
|---|---|
| Base: | tepe |
| Reduplication: | tepetepe |
| Final V Deletion: | tepetep |

Pattern I reduplication may therefore be characterised as involving total reduplication of *CVCV bases.

Pattern II reduplication, partial reduplication, also involves initial CVCV reduplication, as evidenced by the following examples:

| Word | Gloss | Reduplicated |
|---|---|---|
| ngalis | *bite* | ngalingalis |
| sapeng | *answer* | sapesapeng |
| sakone | *force to do* | sakosakone |
| marep | *blink* | maremarep |

Under certain conditions that I will not discuss here, the vowel of the second reduplicated syllable may be reduced to a high vowel, as the following examples illustrate.

| | | |
|---|---|---|
| pirap | *steal* | piripirap |
| lɔnge | *lay across* | lɔngilɔnge |

I am not aware of any data which would support the position that these latter forms involve only CVC reduplication, thus enabling one to treat the high vowel that occurs before the base as an epenthetic vowel. On the contrary, evidence not considered in this paper supports the position that these high vowels may be derived from non-reduced vowels and that all of the examples above involve initial CVCV reduplication.[9]

As the preceding examples illustrate, vowels in the environment CVC___&CVCV... (where & represents the boundary type characteristic of reduplication) may be retained. Under conditions to be specified, however, these vowels may also undergo deletion. As a consequence of such deletion, consonant clusters result, some of which are subject to modification by nasal substitution. The conditions under which these vowels are retained or deleted are examined below.

A vowel in the environment $CVC_1$___&$C_2$VCV... will be retained if (1) $C_1$ and $C_2$ differ in their values for the feature coronal, or (2) $C_1$ and $C_2$ are both non-coronal, but differ in their values for the feature anterior. What happens in the remaining cases is illustrated in the charts below, where the consonants listed down the left side of the chart represent $C_1$ and the consonants listed across the top represent $C_2$. A V is used to indicate that a vowel is retained in the environment specified. If the vowel is deleted, the resulting surface cluster is listed. If, because of co-occurrence restrictions on consonant types within a single morpheme, no example exists, an asterisk (*) is employed.[10] If the missing example is presumed to be due to an accidental gap in the language (or in the data), a dash (-) is listed.

Labials

| | p | pw | m | mw |
|---|---|---|---|---|
| p | mp | * | - | * |
| pw | * | mwpw | * | mwmw |
| m | - | * | mm | * |
| mw | * | - | * | - |

Velars

| | k | ng |
|---|---|---|
| k | ngk | - |
| ng | ngk | ngng |

Coronals

| | s | t | t' | n | l | r |
|---|---|---|---|---|---|---|
| s | ns | - | - | - | V | V |
| t | V | nt | * | V | V | - |
| t' | V | * | nt' | V | V | V |
| n | ns | nt | nt' | nn | ll | - |
| l | ns | nt | nt' | - | ll | - |
| r | ns | nt | nt' | nn | ll | rr |

Examples of the clusters listed in the preceding charts are provided in column three below.

Labials

| Word | Gloss | Reduplicated | Underlying Cluster |
|---|---|---|---|
| pap | *swim* | pampap | pp |
| pwupw | *fall* | pwumwpwupw | pwpw |
| mwopw | *out of breath* | mwomwmwopw | pwmw |
| mem | *sweet* | memmem | mm |

Velars

| | | | |
|---|---|---|---|
| kak | *able* | kangkak | kk |
| kang | *eat* | kangkang | ngk |
| ngong | *bark* | ngongngong | ngng |

Coronals

| | | | |
|---|---|---|---|
| sas | *stagger* | sansas | ss |
| tit | *build a wall* | tintit | tt |
| t'at' | *writhe* | t'ant'at | t't' |
| sinom | *sink in* | sinsinom | ns |
| tune | *tie together* | tuntune | nt |
| t'enek | *hung up* | t'ent'enek | nt' |
| nenek | *commit adultery* | nennenek | nn |
| linenek | *oversexed* | lillinenek | nl |
| sel | *tied* | sensel | ls |
| til | *penetrate* | tintil | lt |
| t'al | *click, tsk* | t'ant'al | lt' |
| lal | *make a sound* | lallal | ll |
| sar | *fade* | sansar | rs |
| tar | *strike, of a fish* | tantar | rt |
| t'ir | *narrowing* | t'int'ir | rt' |
| nur | *contract* | nunnur | rn |
| lirooro | *protective* | lillirooro | rl |
| rer | *tremble* | rerrer | rr |

Examples of potential clusters where a vowel is retained are:

| Word | Gloss | Reduplicated | Flanking Consonants |
|---|---|---|---|
| lus | *jump* | lusulus | s___l |
| rese | *saw* | resirese | s___r |
| set | *artificially ripen breadfruit* | seteset | t___s |
| net | *smell* | netenet | t___n |
| lituii | *serve as female servant* | litilituii | t___l |
| setik | *quick in performing* | set'iset'ik | t'___s |
| net' | *sell* | net'inet' | t'___n |
| let' | *flick* | let'elet' | t'___l |
| rot' | *dark* | rot'orot' | t'___r |

To account for the the preceding data, it is clear that rules of two types are required. First, one or more vowel deletion rules must be posited to explain the occurring consonant clusters. Second, one or more rules of cluster modification are necessary to deal with the fact that the resulting surface clusters are not necessarily identical to the underlying clusters. In the analysis that follows, it is argued that, in fact, nine synchronic rules are

required to account for these data — five rules of vowel deletion and four rules of cluster modification.

In the formulation of these rules, features will be employed only when they permit a more elegant characterisation of a natural class of segments that can be captured through the use of informal notational devices. Also, two ad hoc notational devices will be used in writing these rules. First, following an already common practice, the suprafeature F will be employed to represent all unspecified features. Second, in some rules, two boundary markers will be included within braces to indicate that the rule in question may apply at/across either of these boundaries. The boundary types that play a role in these rules are &, which has previously been identified as the boundary characteristically found between a reduplicated portion of a root and the root itself, and =, which will be used to represent the boundary characteristically occurring between a verb root and directional suffixes.[11] Thus, a rule of the nature A→B/C___{&, =}D specifies that A→B in both the environments C___&D and C___=D. A full discussion of the role of boundaries in PNP, or even of the role they play in these rules, is well beyond the scope of this paper. Data presented in Section 2.2. will make clear, however, the necessity for at least these two boundary types.

The following nine rules are posited to account for the phenomena summarised in the preceding cluster charts. Reference to those charts while examining these rules should prove useful.

1) Vowel Deletion Rule #1

$$V \rightarrow \emptyset / V \begin{bmatrix} +\text{nasal} \\ \alpha\text{ant} \\ \beta\text{cor} \end{bmatrix} \_\_\_\_ \{\overset{\&}{=}\} \begin{bmatrix} +\text{cons} \\ \alpha\text{ant} \\ \beta\text{cor} \end{bmatrix}$$

This rule deletes a vowel in the environment $VC_1\_\_\_\{\overset{\&}{=}\}C_2$, where $C_2$ is a consonant homorganic with the nasal $C_1$.[12] This rule therefore creates the clusters mm, ngk, ngng, ns, nt, nt', and nn, as well as nl, which subsequently becomes ll through the operation of Rule #8.

A V (vowel) must be included as the first segment in the environment of this rule since the first vowel in a word of the shape CV.... may (synchronically) never be deleted or reduced. For example, durative aspect is signalled with morphemes of the shape CVCC or CVVC by reduplicating the initial CV. Note in the following examples of such morphemes that a vowel before & boundary does not delete, even though it is in the correct consonantal environment.

| Word | Gloss | Reduplicated |
|---|---|---|
| mant | *tame* | mamant |
| miik | *suck* | mimiik |

An identical constraint exists on all of the following vowel deletion rules.

2) Vowel Deletion Rule #2

$$V \rightarrow \emptyset / V \begin{bmatrix} +\text{cons} \\ +\text{son} \\ \alpha F \end{bmatrix} \_\_\_\_ \{\overset{\&}{=}\} \begin{bmatrix} +\text{cons} \\ +\text{son} \\ \alpha F \end{bmatrix}$$

This rule deletes a vowel in the environment $VC_1$ ___ $\{^{\&}_{=}\}C_2$, where $C_1$ and $C_2$ are identical sonorants. This rule creates ll and rr clusters, and in the absence of Vowel Deletion Rule #1, would also lead to mm, nn, and ngng. (mwmw occurs only as a derived cluster in reduplication.) This rule could therefore be rewritten to delete a vowel only between potential geminate liquids, but, for purposes of simplicity, it seems preferable to state this rule in its most general form.

3) Vowel Deletion Rule #3

$$V \rightarrow \emptyset / V \begin{bmatrix} +cons \\ -cor \\ \alpha ant \end{bmatrix} \_\_\_\_ \{^{\&}_{=}\} \begin{bmatrix} +cons \\ -cor \\ \alpha ant \end{bmatrix}$$

This rule deletes a vowel in the environment $VC_1$ ___ $\{^{\&}_{=}\}C_2$, where $C_1$ and $C_2$ are either both labials or both velars. When considering reduplicated forms, only the cluster mwmw from *pwmw motivates this rule, since Vowel Deletion Rule #1 and Vowel Deletion Rule #4 will account for all other such clusters. Unlike Rule #4, however, this rule also plays a role in non-reduplicated forms. Its importance in such forms will be further examined at the end of this section and in Section 2.2.

4) Vowel Deletion Rule #4

$$V \rightarrow \emptyset / V \begin{bmatrix} -son \\ \alpha F \end{bmatrix} \_\_\_\_ \; \& \begin{bmatrix} -son \\ \alpha F \end{bmatrix}$$

This rule deletes a vowel in the environment $VC_1$ ___ $\&C_2$, where $C_1$ and $C_2$ are identical obstruents. This rule therefore creates the clusters pp, pwpw, kk, ss, tt, and t't', all of which serve as input to Rule #7, Nasal Substitution.

5) Vowel Deletion Rule #5

$$V \rightarrow \emptyset / V \begin{bmatrix} +son \\ +cor \\ -nas \end{bmatrix} \_\_\_\_ \; \& \; [+cor]$$

This rule deletes a vowel in the environment $VC_1$ ___ $\&C_2$, where $C_1$ is a liquid and $C_2$ is a coronal. This rule creates the clusters ll and rr, as well as the underlying clusters ls, lt, lt', rs, rt, rt', rn, and rl, all of which serve as input to Rule #6.

6) Liquid Assimilation

$$\begin{bmatrix} +son \\ +cor \\ -nas \end{bmatrix} \rightarrow [\alpha F] / \_\_\_ \; \& \begin{bmatrix} +cor \\ \alpha F \end{bmatrix}$$

This rule states that a liquid will completely assimilate to a following coronal across & boundary. This rule applies vacuously to ll and rr clusters, changes ln, lr, rn, and rl to nn, rr, nn, and ll respectively, and changes all homorganic liquid-obstruent clusters to geminate obstruent clusters which serve as input to Rule #7.[13]

7) Nasal Substitution Rule A

$$\begin{bmatrix} -son \\ \alpha F \end{bmatrix} \rightarrow \begin{bmatrix} +son \\ +nas \\ +voice \end{bmatrix} \_\_\_ \; \& \begin{bmatrix} -son \\ \alpha F \end{bmatrix}$$

This rule states that all geminate obstruent clusters that occur with an intervening & boundary will be modified to homorganic nasal-obstruent clusters. Thus, the underlying clusters pp, pwpw, kk, ss, tt, and t't' become mp, mwpw, ngk, ns, nt, and nt' respectively.

8) Nasal Assimilation

$$\begin{bmatrix}+\text{cor}\\+\text{nas}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{nas}\\ \alpha F\end{bmatrix} / \underline{\quad} \; \& \begin{bmatrix}+\text{cor}\\+\text{son}\\-\text{nas}\\ \alpha F\end{bmatrix}$$

This rule states that an n before & boundary will completely assimilate to a following liquid. Thus, in reduplicated forms, nl becomes ll. No reduplicated examples exist of nr becoming rr, but other forms in the language evidence that this change does occur.[14]

9) Nasal Substitution Rule B

$$\begin{bmatrix}+\text{cons}\\-\text{cor}\\ \alpha\text{ant}\\ \beta\text{high}\\ \gamma\text{back}\\ \delta\text{round}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{son}\\+\text{nas}\\+\text{voice}\\ \beta\text{high}\\ \gamma\text{back}\\ \delta\text{round}\end{bmatrix} / \underline{\quad} [-\text{Pause}] \begin{bmatrix}+\text{cons}\\-\text{cor}\\ \alpha\text{ant}\\ \beta\text{high}\\ \gamma\text{back}\\ \delta\text{round}\end{bmatrix}$$

This rule states that if two labial or two velar consonants come together in the flow of speech (that is, no pause intervenes), the first consonant will become a nasal that copies the features of velarisation (high, back, round) of the second. Within the data previously examined, only the modification of the cluster pwmw to mwmw illustrates the operation of this rule. But this rule is well supported by forms in which reduplication is not involved. The operation of this nasal substitution rule outside of reduplication will be examined in Section 2.2.

The preceding rules will account for all of the surface clusters listed in the charts at the beginning of this section. Admittedly, the fact that nine rules are required to explain these data suggest an inelegant or inaccurate solution. Especially suspicious is the necessity for five rules of vowel deletion. It is by no means obvious, however, how the number of such deletion rules could be reduced. A solution wherein all vowels are deleted before reduplication boundary, and subsequently vowels are inserted to break up impermissable clusters, is ruled out by the fact that the vowels that occur between such potential clusters are not predictable; they are underlying vowels. A solution in which two or more of these deletion rules are collapsed into a single rule also suggests itself, but, in such a solution, one must take into account the fact that the vowel deletion rules previously listed are of two types — those that apply both at & and = boundary, and those that apply only at & boundary. Thus, if the number of vowel deletion rules is to be reduced, one must find a way to combine rules 1 and/or 2 and/or 3 as one set and rules 4 and 5 as another set. The collapsing of rules within these constraints is not possible. It is, of course, possible to restate these vowel deletion rules, and to reorder them, but none of these alternant solutions is better motivated, nor does any come as close as the existing solution to capturing what must have happened historically in the language.[15] Given the current status of research on PNP, nine rules are required.

## 2.2. Nasal substitution outside of reduplication

Of the two nasal substitution rules cited above, only Rule B applies to clusters outside of reduplication. Rule B states that whenever two labial or two velar consonants come together in speech, therefore no pause intervenes, the following results obtain:

| | Labials | | | | Velars | | |
|---|---|---|---|---|---|---|---|
| | p | pw | m | mw | | k | ng |
| p | mp | mwpw | mm | mwmw | k | ngk | ngng |
| pw | mp | mwpw | mm | mwmw | ng | ngk | ngng |
| m | mp | mwpw | mm | mwmw | | | |
| mw | mp | mwpw | mm | mwmw | | | |

Examples of the application of this rule in polymorphemic words of various types are presented below. The clusters resulting from this rule are underlined in the first column. In the second column, the morphemes of which these words consist are listed as they would occur after the application of all rules except Nasal Substitution Rule B. The dash in these forms simply indicates the presence of a boundary marker which, for the purposes of these examples, may remain unspecified. Eng. glosses are provided in the third column.

| Word | Consisting of the Morphemes | Gloss |
|---|---|---|
| lompeseng | lop-peseng | *cut apart* |
| sarempene | sarep-pene | *scrape together* |
| tɔtɔɔngki | tɔtɔɔk-ki | *work with* |
| isingki | isik-ki | *burn with* |
| keemmeirkelik | keep-meir-kelik | *yam variety* |
| emwpwɔt'ol | ep-pwɔt'ol | *a game* |
| keemwmwot' | keep-mwot' | *yam variety* |
| sampaa | sapw-paa | *world, earth* |
| limwpwɔt' | lim-pwɔt' | *five (oblong things)* |
| limwmwut' | lim-mwut' | *five (piles)* |

Note that the last five examples illustrate regressive assimilation to the features of velarisation.

At normal conversational speed, this rule applies even across word boundary, as illustrated by the following examples:

Sentence: E kalap paan soupisek.
Pronounced: /e kalam paan soupisek/
Gloss: *He'll always be busy.*

Sentence: E kalap men meir.
Pronounced: /e kalam men meir/
Gloss: *He's always sleepy.*

Sentence: E saik kengwini.
Pronounced: /e saing kengwini/
Gloss: *He hasn't yet taken medicine.*

Sentence: E saik nget.
Pronounced: /e saing nget/
Gloss: *He's not yet out of breath.*

Sentence: Soulik kin soupisek.
Pronounced: /souling kin soupisek/
Gloss: *Soulik is (habitually) busy.*

The last example (where Soulik is a noun phrase and kin soupisek is a verb phrase) illustrates that this rule also applies to segments belonging to different syntactic phrases.[16] The constraint on Nasal Substitution Rule B, then, is that it applies to segments within the same phonological phrase; that is, to segments which are not separated by a pause.

Garvin's constraint on nasal substitution, that it applies to segments within the same phrase (assuming that he was referring to phonological phrases) is accurate for homorganic labial and velar clusters. But, coronal consonants do not behave in a parallel manner. Coronal consonants undergo nasal substitution in reduplicated forms only. The nasal substitution rule affecting coronal consonants, Nasal Substitution Rule A, never applies across word boundary, as the following examples illustrate.

Sentence: E ekis suwet.
Pronounced: /e ekis suwet/
But Never: */e ekin suwet/
Gloss: *It's kind of bad.*

Sentence: Ke meit tangaanga!
Pronounced: /ke meit tangaanga/
But Never: */ke mein tangaanga/
Gloss: *Aren't you lazy!*

Coronal consonants also do not undergo nasal substitution in polymorphemic words, typically because coronal clusters do not arise in such words. Therefore, Vowel Deletion Rules 4 and 5 are constrained so as to apply at & boundary only. In words involving other boundaries, these rules do not apply and coronal clusters do not result, as illustrated by the following examples.

| Word | Consisting of the Morphemes | Gloss |
|---|---|---|
| isisel | isi+sel | *seven (ropes)* |
| isisop | isi+sop | *seven (stalks)* |
| palisal- | pali+sal- | *side exposed to* |
| weitita | weiti=ta | *proceed upward* |
| pɔteti | pɔte=ti | *plant downward* |
| lusisang | lusi=sang | *jump from* |

In these examples, the final vowel of the base is retained before + and = boundary. However, even where an enclitic boundary occurs, before which base final vowels delete, nasal substitution does not apply. Either a copy vowel is inserted to break up the coronal cluster or the cluster occurs without further modification. Note the following examples:

| | | |
|---|---|---|
| mwemeit'et'e | mwemweit'#t'e | *just visiting* |
| massuwet | mas#suwet | *ugly* |

The example Garvin cited of mwoonti *sit down* from mwoot *sit* and -ti *down* would appear to violate the claim that coronal consonants do not participate in nasal substitution except in reduplication, but in fact this form is anomalous. This same verb followed by the directional suffix -ta *upward*, for example, results in the form mwootata. One explanation for the occurrence of the form mwoonti rather than mwootiti (which some speakers also accept) is that this verb so often occurs with this particular suffix that the boundary normally present between verbs and directional suffixes (= boundary) was replaced by the tighter boundary & and the correct environment for Nasal Substitution Rule A arose. It is also possible that this form simply underwent lexicalisation early

in the history of PNP and that diachronically Nasal Substitution Rule A operated morpheme-internally as well as across & and possibly other boundaries. Some support for this latter position is offered in the section that follows.

## 3. DIACHRONIC APPLICATIONS OF NASAL SUBSTITUTION RULE A

Synchronically in PNP, homorganic nasal obstruent clusters occur not only in polymorphemic words, but morpheme-internally as well. Morpheme-internal examples of such clusters follow.

| | Initially | Medially | Finally |
|---|---|---|---|
| mp | (i)mpe<br>*next to (it)* | tempel<br>*kava pounding rhythm* | emp<br>*coconut crab* |
| mwpw | (u)mwpwul<br>*flame* | semwpwe<br>*turn windward* | (u)mwpwɔmwpw<br>*low hill* |
| nt | (i)ntil<br>*torch fish* | manta<br>*next day* | mant<br>*tame* |
| ns | (i)nsar<br>*snare* | kounsup<br>*frowning* | kens<br>*yaws* |
| nt' | (i)nt'a<br>*blood* | eent'a<br>*fish sp.* | kent'<br>*urine* |
| ngk | (i)ngket<br>*thatch* | lingkiri<br>*oyster* | engk<br>*landslide* |

The vowel enclosed in parentheses before initial occurrences of such clusters is a predictable vowel. Further comments on the origins of this vowel will be presented later in this section.

Considerable evidence exists to support the position that at least some of these clusters arose from earlier geminate obstruent sequences that underwent Nasal Substitution Rule A, or some historical antecedent of this rule. Considering PNP data alone, this position is weakly supported by the nature of the distribution of geminate consonants morpheme-internally.

| Occurring | Non-Occurring |
|---|---|
| mm | *pp |
| mwmw | *pwpw |
| nn | *tt |
| ll | *ss |
| rr | *t't' |
| ngng | *kk |

Note that whereas all sonorants occur geminate morpheme-internally, obstruents never do.[17] The skewed distribution of these segments could obviously be explained by assuming that all geminate obstruent clusters were modified by nasal substitution.

The most perusasive evidence in support of this position, however, comes from external comparisons with other MC languages. Compare, for example, the previously cited PNP forms containing initial nasal-obstruent clusters with their cognates in Trukic (TK). (Lagoon Trukese (TRK) forms are employed except where noted.)

| PNP | TK |
|---|---|
| (i)mpe<br>*next to (it)* | ppa-<br>*next to --* |
| (u)mwpwul<br>*flame* | pwpwun<br>*flame* |
| (i)ntil<br>*torch fish* | ttúl- (Saipan Carolinian)<br>*fish with a light* |
| (i)nsar<br>*snare* | ssar<br>*snare* |
| (i)nt'a<br>*blood* | čča<br>*blood* |
| (i)ngket<br>*thatch* | kket<br>*thatch* |

In these examples, the nasal-obstruent clusters of PNP systematically correspond to the geminate obstruent clusters of TK.

Since geminate consonants do not occur initially or in any other position in Proto-Oceanic (POC) reconstructions, the presence of such clusters constitutes a problem in understanding the history of these languages. Goodenough (1963) recognised this problem and argued that morpheme-internal geminates in TRK developed as a consequence of vowel deletion in earlier morphologically complex forms. TRK morphs with initial geminates, he stated "reflect older forms with classic first syllable reduplication" (Goodenough 1963:78), as evidenced by cognates in Kiribati (KIR) (Gilbertese).

| KIR | TRK |
|---|---|
| raraa<br>*blood* | čča<br>*blood* |
| kakang<br>*sharp* | kken<br>*sharp* |

Such reduplication is not evidenced in POC (*daRa(?) *blood*),[18] but it does apparently date back as far as Proto-Micronesian (PMC *t'at'aa *blood* and *kakangi *sharp*).[19] What function this pattern of reduplication served is not entirely clear.

It is also not entirely clear how a PNP form such as (i)nt'a *blood* developed from PMC *t'at'aa, but one possible scenario follows.

| | PMC | *t'at'aa |
|---|---|---|
| 1) | Metathesis | at't'aa |
| 2) | Nasal Substitution | ant'aa |
| 3) | Vowel Reduction | int'aa |
| 4) | Final Vowel Deletion | int'a |
| 5) | Reduced Vowel Deletion (Optional) | (i)nt'a |

Comments concerning the motivation for constructing this particular scenario follow.

That a rule of metathesis might have existed diachronically in PNP is supported by evidence from Marshallese (MRS). The developments affecting potential initial geminates in the Ratak and Rālik dialects of this language are particularly suggestive of what might have happened in PNP.[20]

| Ratak | Rālik | Gloss |
|---|---|---|
| kẹkẹn | yẹkkẹn | *invent* |
| liliw | yilliw | *angry* |

The Ratak forms are the result of earlier CV- reduplication, where the vowel between the potential geminates is not deleted. In the Rālik dialect, however, as possibly in PNP, metathesis has taken place (along with the development of a prothetic y) and geminates result.

The diachronic rule governing nasal substitution was apparently identical to Nasal Substitution Rule A, presented in the last section, with one important difference. Whereas the synchronic rule of nasal substitution must be constrained so as to operate only across & boundary, the diachronic rule apparently also operated morpheme-internally. Considering only forms involving earlier initial CV- reduplication, this position is tenuous, but can be argued. Therefore, since a form such as *t'at'aa *blood* was already inherently reduplicated in PMC, it does not seem unlikely that this form at some point relatively early in the history of these languages underwent lexicalisation, so that no internal boundary remained. At least, it does not seem unlikely that this boundary information was lost by the time nasal substitution applied, which was apparently no earlier than Proto-Ponapeic (PPP). Nasal substitution is not found outside this subgroup. Since the time depth between PMC and PPP is unknown, however, this argument is weak. Further justification for this position is necessary and will be provided in the discussion of the origin of medial and final nasal-obstruent clusters.

To account for the synchronic shape of a word such as (i)nt'a *blood*, a diachronic rule of vowel reduction is also required. The effect of this rule was to reduce metathesised vowels to either u or i, depending upon the rounding of the following segment.[21] Therefore, these vowels were reduced to u before (1) clusters of velarised labial consonants, and (2) clusters of other consonants followed by a round vowel, thus leading to the rounding of the preceding segments. Elsewhere, these vowels reduced to i. This is the synchronic situation, as illustrated by the following examples.

| | |
|---|---|
| Initial u | |
| umwpwer | *twin* |
| ungkopw | *crab sp.* |
| Initial i | |
| impe | *next to (it)* |
| inta | *say* |
| int'a | *blood* |
| ingkapwan | *a while ago* |

The last two developments, Final Vowel Deletion and the optional rule of Reduced Vowel Deletion, are well attested synchronic rules of PNP.

One might, of course, construct alternate scenarios that would equally well account for synchronic surface forms like (i)nt'a. I am not certain, for example, whether Nasal Substitution was a later or earlier development than Vowel Reduction. I am confident, however, that the optional, predictable vowels that occur in such forms represent vowels that, from a historical perspective, are in the process of being lost, rather than added. In earlier analyses of PNP, I had in fact taken the opposite position.[22] Because these vowels are optional and predictable, I assumed that they originated as prothetic vowels. While such an analysis might be possible in a synchronic grammar of

PNP, it is now clear that forms such as (i)nt'a must have developed from an earlier int'a. As evidence for this position, consider the fact that among all the speech communities that make up Ponapeic (PP), it is only in PNP that these vowels are optional. In Ngatikese (NGK), for example, the initial vowels before nasal-obstruent clusters may never be deleted, as illustrated by the following examples.[23]

| PNP | NGK | |
|---|---|---|
| (i)mpaai | impaai | *submissive* |
| (u)mwpwɔs | imwpwos | *boil* |

Note, further, that in NGK the vowel that precedes these clusters is always i, regardless of the rounding of the following segment. In light of this observation, the following forms are particularly interesting.

| PNP | NGK | |
|---|---|---|
| (u)mwpwel | imwpwel | *crab sp.* |
| (u)mwpwel | umwpwel | *earth oven filled with food* |

While the first NGK example has an initial i, the second does not. It is not the NGK form which is aberrant, however. It is the PNP form. In both PNP and NGK, the word umwpwel is, historically at least, a compound consisting of the two morphemes umw *earth oven* and pwel *earth*.[24] In PNP, therefore, the first vowel of the word umwpwel is being treated precisely like the metathesised reduced vowels of earlier reduplicated forms and is consequently subject to deletion.

Further support for the position that these vowels are being lost is provided by a comparison of the following forms in PNP and Mokilese (MOK) (Harrison 1977).

| PNP | MOK | |
|---|---|---|
| mmat' | immas | *ripe* |
| mwmwus | umwmwuj | *vomit* |

The geminate sonorants that occur in these forms in both languages also apparently arose as a consequence of initial CV- reduplication, metathesis, and vowel reduction.[25] In MOK, the resulting initial high vowels are retained in these forms, but in PNP they are completely lost. Thus, the rule optionally deleting initial vowels before homorganic nasal-obstruent clusters in PNP appears to be an extension of a rule which historically deleted such vowels before geminate sonorants. There thus appears to be little doubt that the optional initial vowels that we have been examining represent vowels in the process of being lost rather than added. The importance of this in explaining the motivation for nasal substitution will be examined in the next section of this paper.

So far as I am aware, except for the aberrant form (u)mwpwel, all surface initial occurrences of nasal-obstruent clusters derive historically from underlying geminate obstruents, where the first obstruent underwent nasal substitution. The origin of morpheme-internal nasal-obstruent clusters in other positions, however, is not so well understood, primarily due to the difficulty of finding non-PP cognates with these forms. Apparently, though, these clusters arose in two distinct ways.

First, non-initial nasal-obstruent clusters arose morpheme-internally as a result of the deletion of a vowel that was preceded by a nasal and followed by a consonant homorganic with that nasal. Vowel loss in this environment

occurred in earlier morphologically complex forms, as illustrated by the following example.

| | | |
|---|---|---|
| menseng | From: | mani+sangi |
| *morning* | | *animal+cry* |

Vowels were also lost in this environment in monomorphemic words, however, as illustrated by the next example.

| | | |
|---|---|---|
| mant | From: POC | *manasa |
| *tame* | | *tame* |

The rule that governed this vowel deletion was the progenitor of synchronic Vowel Deletion Rule #1, cited in the preceding section. For reasons that I will not explore in this paper, the synchronic version of this rule must be constrained so that it will operate only at & and = boundaries, but, as the form mant illustrates, this rule did operate morpheme-internally diachronocially.

Second, non-initial nasal-obstruent clusters arose morpheme-internally as a consequence of nasal substitution. Such clusters occur in forms involving a now-fossilised pattern of -CVCV final reduplication, as the following example illustrates.[26]

| | | |
|---|---|---|
| opampap | From: | *opapa&papa |
| *humble* | | |

Vowel Deletion Rule #4 and Nasal Substitution Rule A, as synchronically stated, will account for the preceding form, but some evidence exists that both of these rules also operated morpheme-internally. Consider, for example, the word for 'coconut crab' in the following languages.

| | |
|---|---|
| PNP: | emp |
| MOK: | opup |
| Woleaian (WOL): | yaff |

Further research will be required to determine precisely how the synchronic vowel deletion and cluster modification rules operated diachronically. It seems likely, however, that all of these rules played a role in the origin of morpheme-internal clusters in modern PNP.

## 4. THE MOTIVATION FOR NASAL SUBSTITUTION

The presence, either diachronically or synchronically, of nasal substitution rules is one of the defining characteristics of PP languages. No other nuclear MC language exhibits this phenomenon, nor, so far as I am aware, does any other OC language. Nasal substitution is by no means a common phonological development. Thus, why such rules should have developed in PP is a question of both historical and theoretical interest.

Fischer (1965:1496) takes up this question and argues that nasal substitution in PNP is stylistically motivated. He notes that, across word boundary, TRK and PNP treat homorganic consonant clusters in antithetical ways. "In Trukese, the preference is for clusters of two identical stops (articulated as a single long stop), while in Ponapean the preference is for clusters of nasal plus stop." The Sandhi rules presumably leading to these clusters, he notes, are as follows (where **N** = nasal, S = stop, and subscript **a** = point of articulation).

| | Underlying Forms | Result of Sandhi |
|---|---|---|
| Trukese | $...N_a\ S_a...$ | $...S_a\ S_a...$ |
| | $...S_a\ S_a...$ | $...S_a\ S_a...$ |
| Ponapean | $...N_a\ S_a...$ | $...N_a\ S_a...$ |
| | $...S_a\ S_a...$ | $...N_a\ S_a...$ |

Based on the cultural attitudes toward speech as they are influenced by the social structure of these speech communities, and on the occasions on which he heard the sandhi rules apply or fail to apply in PNP, Fischer (1965:1500) hypothesises that

> there is a common expressive significance for each of the two types of consonant clusters in the two languages. Specifically, in both languages the double stops seem to have the value of abruptness and freedom of emotional expression, while the clusters of nasal plus stop seem to have the value of restraint, politeness, and gentleness.

Ponapeans, who value restraint and the avoidance of overt expressions of aggression in speech, prefer nasal-stop clusters because of their symbolic value. Trukese, who are more aggressive in their speech behaviour, prefer double stops. He further suggests that the patterns displayed by TRK and PNP might be explained in terms of a "potentially universal symbolism" (p.1500). Double stops involve a more forceful vocal gesture, and may thus be expressive of aggression, while nasal-stop clusters are less forceful, and may thus be expressive of restraint.

I do not disagree that phonetic symbolism plays a role in language, but I think Fischer's position, that such symbolism motivated nasal substitution, fails in two ways. First, Fischer's account of the rules which govern nasal substitution follows Garvin's analysis, and is thus wrong. As noted in the second section of this paper, coronal stop clusters do regularly occur across word boundary and, within certain prosodic configurations, in other environments as well. It is not simply the case, therefore, that geminate stops are avoided in PNP; their occurrence is governed by a complex series of phonological rules. Second, I suspect that while phonetic symbolism may play a role in determining the shape of some morphemes, it does not normally, if ever, dictate regular sound change.[27] Labov has noted, of course, that social and stylistic factors play a role in sound change, but these factors serve to influence the utilisation or non-utilisation of otherwise phonetically or functionally motivated processes. The motivation for nasal substitution can, I believe, be explained in these terms.

One obvious motivation for nasal substitution is the functional role it plays in limiting the number of optimal consonant cluster types in PNP. As demonstrated in the second section of this paper, nasal substitution rules interact with *a* complex series of other rules as part of a conspiracy to reduce 144 potential consonant cluster types to 12 optimal ones. Thus, nasal substitution is motivated in part by the role it plays within the phonological system of PNP.

Nasal substitution would also appear to be well-motivated on perceptual grounds. Voiceless geminate obstruents are difficult to perceive, especially when they are in initial or final position.[28] By lowering the velum and adding voicing to the first obstruent — the changes involved in nasal substitution — this perceptual problem is alleviated. Many languages, of course, do tolerate

voiceless geminate obstruents, and some languages (MOK for example) even have rules which create such clusters.[29] On the whole, though, rules which simplify such clusters appear to be far more common, probably for the reasons cited above.

The lowering of the velum involved in nasal substitution also suggests a physiological motivation for this phenomenon. Sheldon Harrison first noted that the conversion of an obstruent to a nasal in this environment

> can be viewed as a response to the heightened pressure inherent in geminate obstruents. This pressure can be reduced by lowering the velum to allow some air to escape through the nasal cavity, thereby destroying the obstruent articulation.[30]

Another motivation for nasal substitution rests on the claim I wish to advance, that this phenomenon is not the result of dissimilation, as Garvin proposes, but rather is a consequence of weakening. As Johnson (1973:52) has insightfully noted, "What appears...as a dissimilatory state...may not be due to a dissimilatory process." Johnson, along with Foley (1972), takes the position that many seemingly dissimilatory processes involving contiguous segments may be better explained in terms of weakening. It is therefore of significance that in PNP, both synchronically and diachronically, nasal substitution affects consonants in syllable-final position. The two rules of nasal substitution operate in the following environments (where $ represents syllable boundary):

| Initial | Medial | Final |
|---|---|---|
| #VC$CV... | ...VC$CV... | ...VC$CV# |

Hooper (1976:196) and others have argued that certain environments are particularly conducive to weakening. One such environment for consonants is syllable-final position, the environment described above.

The terms 'weakening' and 'strengthening' are commonly employed in discussions of sound change, but as Sommerstein (1977:228) has noted, "their definitions have tended to remain intuitive rather than explicit..." Weakening processes for consonants, however, are typically those which result in less obstruction and/or increased sonority, with the extreme case of weakening being deletion. Based upon the extreme case of deletion, an explicit definition of weakening has been attempted by Hyman (1975:165) who states "a segment X is said to be weaker than a segment Y if Y goes through an X stage on its way to zero". Hyman's definition very nicely supports the position that nasal substitution in PNP is, in fact, weakening, as evidenced by data from Pingelapese (PNG), a PP language bordering on mutual intelligibility with PNP. In PNG, the word for *blood*, for example, is iisa (cf. PNP (i)nt'a).[31] Considerable evidence exists to support the position that this form developed from an earlier issa that underwent nasal substitution, resulting in insa. Subsequently, the n was deleted and the preceding vowel was lengthened to maintain the original number of moras in the word. Thus, following Hyman's definition, it can be argued that nasals are weaker than obstruents, since, as PNG evidences, obstruents go through a nasal stage on their way to zero.

The motivation for the weakening of consonants in syllable-final position, and thus a motivation for nasal substitution, is almost certainly related to the fact that the single universal syllable type is the open syllable - CV. Within OC, and PNP, this is the dominant syllable type. Substituting a nasal for a syllable-final obstruent would appear to be one way of opening up the syllable, as evidenced by the many languages in the world that permit only open syllables or syllables ending in a nasal. Thus, nasal substitution may represent

an attempt to restore the optimal pattern of syllables in PNP, sequences of open syllables, that was violated by earlier vowel deletion rules.

It is not possible, nor is it even desirable, to single out one of the motivations I have suggested as being the 'correct' or the 'major' motivation for nasal substitution. Quite likely, all of these considerations played a role in the development of this phenomenon. The problem posed by Fischer - why a related language like TRK did not develop nasal substitution, but in fact developed rules creating geminate obstruents - of course remains. It is not the task of this paper to answer that question, but it seems likely that in TRK sufficient numbers of surface geminate obstruents occurred that a change was effected in the preferred syllable type. Further research into this question is required.

## NOTES

1. I wish to thank Byron W. Bender, Iovanna Condax, John L. Fischer, Ward Goodenough, Jimmy Harris, Irwin Howard, Jeff Marck, and David Stampe for discussing with me some of the ideas expressed in an earlier draft of this paper. The responsibility for the final form of this work is mine alone.

2. Ponapean (PNP) is spoken by approximately 16,000 residents of the island of Ponape in the Eastern Caroline Islands.

3. Included in nuclear Micronesian are Kiribati (KIR) (formerly Gilbertese), Marshallese (MRS), Kosraean (formerly Kusaiean) (KSR), the Trukic (TK) languages, the Ponapeic (PP) languages, and possibly Nauruan. The PP subgroup includes PNP, Ngatikese (NGK), Mokilese (MOK), and Pingelapese (PNG).

4. The consonant represented as ts in this chart is labelled post-alveolar. The precise position of articulation of this consonant, however, is in doubt. So far as I have been able to determine, it is produced by placing the blade of the tongue against or slightly behind the alveolar ridge, while the sides of the tongue are in contact with the upper gums approximately as far back as the palatal region. Further comments on the segmental phonemes of PNP are presented in Rehg 1973.

5. I have not yet determined precisely how this process should be formalised. It operates, however, in at least the following environments (where C represents a coronal consonant and $ represents a syllable boundary): nC, nVC, and nV$CV$.

6. The insertion of epenthetic vowels is briefly commented on in Section 2.2. It should also be noted that while geminate obstruents never occur morpheme-internally in native vocabulary, they do occur in some loan words from Jp. (e.g. nappa *Chinese cabbage*). They also arise across morpheme boundaries in particular prosodic configurations.

7. PNP also evidences several fossilised patterns of reduplication, including final -CVCV reduplication. This latter pattern is briefly commented on in Section 3.

8. Only high and mid vowels surface in the environment CVC___&CVC, where V is a non-low vowel, due to a rule which raises low vowels in this and similar environments.

9. In certain segmental and/or prosodic environments, full vowels are reduced to high vowels. High vowels in PNP function as 'minimal' vowels and, in these same environments, are subject to deletion in rapid speech.

10. These co-occurrence restrictions are discussed in Rehg 1973, Section 2.3.

11. A third boundary not dealt with in this paper is the boundary characteristically found between nouns and possessive suffixes.

12. For the purposes of this rule, t' is treated as +anterior. However, if t' is in fact articulated far enough behind the alveolar ridge that it must be considered -anterior, then this rule will not work. In this case, one could either formulate two rules in place of Vowel Deletion Rule #1 — one which would operate on coronals and one which would operate on non-coronals — or one could allow the process described in note 5 to operate before this rule.

13. In MOK, a PP language in which nasal substitution is no longer productive, homorganic liquid-obstruent clusters surface as geminate obstruents. In MOK, as in PNP, this rule operates only in reduplicated forms. See Harrison (1976:45).

14. This rule operates optionally in other environments as well. Therefore, nanleng *heaven* (from nan *in* and laang *sky*) may also be pronounced nalleng. Similarly, nan Ruk *in Truk* may in rapid speech be pronounced nar Ruk.

15. Actually, the overlapping involved in these rules, at least historically, is not surprising. For example, because Rule 1 accounts for nasal-obstruent clusters as well as nasal-sonorant clusters, and in fact deletes vowels between all geminate sonorants except ll and rr, a new rule, Rule 2, arises which takes care of these exceptions. Synchronically, these rules have not progressed to the point where vowels simply delete before & boundary, but one can speculate this is the direction in which the language is heading. In fact, MOK apparently has moved rather far along toward just such a treatment of these vowels.

16. Patterns of intonation in PNP, however, are determined by syntactic constituents. Thus, suprasegmental and segmental phenomena are governed by phrases of different types.

17. Except, as noted in note 6, in some Jp. loan words.

18. POC reconstructions are from Grace (1969).

19. PMC reconstructions are from Marck (1977).

20. From Abo et al. (1976) and Karen Kaeo (personal communication).

21. The notion of 'reduction' is appropriate here since non-high vowels are changed to high vowels, the minimal vowels in PNP. See note 9.

22. I stated this position in a number of unpublished papers, and it is implicit in Rehg (1973).

23. NGK is mutually intelligible with PNP. It is spoken on the atoll of Ngatik, located approximately 90 miles south-west of Ponape.

24. This form is also interesting from the point of view of reconstructing the material culture of Ponape. While the contemporary Ponapean oven is built on the surface of the ground, and is never covered with earth, this form suggests an earlier technique of building such ovens underground, as is the more common practice in the Pacific.

25. Vowel reduction applies vacuously in the case of *umwmwus.

26. See Harrison (1973) for a discussion of final reduplication in MC languages.

27. For example, surface-initial (u)mwpw clusters in PNP do seem to have a phonoaesthetic function. Morphemes beginning with this sequence of sounds typically involve the semantic notion of 'roundness', as illustrated by the forms (u)mwpwe *curve*, (u)mwpwei *ball*, (u)mwpwek *bud of a flower*, (u)mwpwel *blister*, (u)mwpwet *to blister*, (u)mwpwi *a drop*, (u)mwpwokos *humpback*, (u)mwpwos *boil*, (u)mwpwɔmwpw *low hill*, (u)mwpwun *barnacle*, and so on. Final əmp clusters in English seem to have a similar value, as evidenced by *hump*, *lump*, *rump*, *clump*, *dump*, etc.

28. But it is possible to perceive even geminate stops in these positions. The articulation may be more fortis, and slight differences in the colouring of the adjacent vowel may occur. Also, there may be visual clues, such as the shape of the speaker's mouth.

29. See note 13.

30. Harrison (1983:359). This motivation for nasal substitution was also suggested to me independently by Ward Goodenough and Jimmy Harris.

31. Elaine Good, personal communication.

## BIBLIOGRAPHY

ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DeBRUM

1976 *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.

BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 426-465. The Hague: Mouton.

FISCHER, John L.

1965 The stylistic significance of consonantal sandhi in Trukese and Ponapean. *American Anthropologist* 67/6(1):1495-1502.

FOLEY, James

1972 Rule precursors and phonological change by meta-rule. In Robert P. Stockwell and Ronald K.S. Macauley, eds *Linguistic change and generative theory*, 96-100. Bloomington: Indiana University Press.

GARVIN, Paul L.

1962 Ponapean morphophonemics. *Phonetica* 8/1-3:115-127.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The proceedings of the Ninth Pacific Science Congress, 1957*, vol.3:77-86.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

HARRISON, Sheldon P.

1973 Reduplication in Micronesian languages. *Oceanic Linguistics* 12:407-454.

1983 Segmental quantity in Mokilese: a synchronic and diachronic study. In Byron W. Bender, ed. *Micronesian Studies,* pp.375-401. *PL,* C-80.

HARRISON, Sheldon P. and Salich ALBERT

1976 *Mokilese reference grammar*. Honolulu: University Press of Hawaii.

1977 *Mokilese-English dictionary*. Honolulu: University Press of Hawaii.

HOOPER, Joan B.

1976 *An introduction to natural generative phonology*. New York: Academic Press.

HYMAN, Larry M.

1975 *Phonology: theory and analysis*. New York: Holt, Rinehart and Winston.

JOHNSON, Lawrence

1973 Dissimilation as a natural process in phonology. *Stanford Occasional Papers in Linguistics* 3:45-56.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

REHG, Kenneth L.

1973 On the history of Ponapean phonology. *Working Papers in Linguistics, University of Hawaii* 5/8:17-55. See also Rehg 1983.

1981 *Ponapean reference grammar*. Honolulu: University Press of Hawaii.

SOMMERSTEIN, Alan H.

1977 *Modern phonology*. London: Arnold; Baltimore: University Park Press.

WELLEY, Weldis and Elaine GOOD

1976 A short grammar of Pingilapese. MS.

Rehg, K. "Nasal substitution rules in Ponapean". In Bender, B.W. editor, *Studies in Micronesian linguistics*.
**C-80:317-337**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.317*

# MOKILESE REFLEXES OF PROTO-OCEANIC

Sheldon P. Harrison

## 1. SOME REMARKS ON MOKILESE (MOK) PHONOLOGY[1]

Comparisons between MOK and Proto-Oceanic (POC) are made on the basis of the following taxonomic phonemic inventory for MOK:

| Consonants | Labials | Dentals | Palatals | Velars | Labio-Velars |
|---|---|---|---|---|---|
| Stops | p | t | j | k | pw |
| Nasals | m | n | | ŋ | mw |
| | | s | | | |
| | | l | | | |
| | | r | | | |

| Vowels | | | |
|---|---|---|---|
| | i | u | Long vowels are |
| | e | o | represented as geminate. |
| | ɛ | ɔ | |
| | a | | |

| Glides | | |
|---|---|---|
| | y | w |

Post-vocalic glides are occasionally represented as i and u. This is not intended to reflect any phonetic difference, but is merely an inconsistency in the transcription.

The sound represented by j is phonetically a prepalatal stop or affricate.

As regards the vowels, all occur long, although there is some question as to their distribution in final position. (uu, ee and ɛɛ are not attested in final position.) The vowel e is of low frequency. The vowels o, ɔ, and u have obviously centralised allophones in certain consonantal environments, the details of which are not yet completely clear. For example:

| | |
|---|---|
| pwo [pwo] | *sterile* |
| po [pə] | *deck* |
| pwɔ [pwɔ] | *fern* |
| pɔ [pʌ] | *arm* |
| pwu [pwu] | *betel nut* |
| pu [pɨ] | *bent* |

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 339-373. *Pacific Linguistics*, C-80, 1984.

In addition, initial u sometimes varies with wi. Thus, uk *blow* is often homophonous with wik *change*, while upw *young coconut* has no alternate **wipw.[2] Long back rounded vowels seem to preserve their backness in all environments.

The glides appear initially, medially, finally, and preconsonantally. Their phonemic status remains undecided.

| | |
|---|---|
| pɛyn | *female* |
| ya | *who? which?* |
| keyɔw | *first* |
| pɛy | *fight* |
| tawn | *fill* |
| woy | *turtle* |
| jowi | *conch* |
| ɛw | *one*, general classifier |

## 2. MORPHEME STRUCTURE OF MOK

As has been observed for Trukese (TRK) (Dyen:1949:421) and other 'nuclear' Micronesian (MC) languages, MOK does not normally reflect POC final vowels.

*laŋo *(a) fly* > lɔŋ *(a) fly*

Thus, all monomorphemic (unaffixed and noncompound) reflexes of POC vowel-final items should be consonant-final. In cases where this is not true we must assume either loss of $C_2$ in a POC *$C_1V_1C_2V_2$# structure, subsequent addition of a final syllable, or some aberrant development.[3]

The identity of lost final vowels can be inferred in many cases from the quality of synchronic reflexes of historically non-final vowels. Thus

| | | | | |
|---|---|---|---|---|
| mar | *preserved breadfruit* | < | *mada | *fermented* |
| jɔŋ | *weep* | < | *taŋi(s) | *weep* |

That is, final non-low vowels raised a preceeding *a, while low vowels had no similar effect. (This will be treated more fully in Section 4.1.).

I am also led to believe that synchronic reflexes of POC final vowels appear on the surface in certain paradigmatic alternations; in particular, the possessive forms of inalienable nouns and the transitive forms of verbs. I would like to suggest that these reflexes appear in some cases because an older syllabic suffix has been lost. For example:

| | | | | |
|---|---|---|---|---|
| *ansa(n) | *name* | > | at | *name* |
| | | > | ɔtɔ | *his name* |
| *nsoka | *stab* | > | dok ]$_{int}$ | *stab* |
| | | > | dɔkɔ ]$_{tr}$ | |

This is not to say that these paradigms are a perfect reflection of the POC final vowels. For example, winin *his hair* < *ʔunap *fish scale*, where *a ⊁ i in this environment (a doublet winɛ *its feathers, scales* from the same source shows a regular reflex). My data contain no comparable examples for transitive verbs, but I do not doubt that they could be found.

Synchronic facts about MOK lead me to conclude that the productive pattern for the creation of inalienable nouns is the i-stem pattern (for example, nuusin *news of him* < Eng. news), and that for transitive verbs the addition of the suffix -i (aini *iron it* < Eng. iron). Thus, whereas many verbs and nouns come by their i-stem vowels 'honestly' (juki *pound it* < *tuku *pound* and kilin *its bark* < *kuli(t) *skin, bark*), the mere fact that a given lexical item is an i-stem is not necessarily evidence that the historical vowel was in fact one that yielded i. This does not seem to be the situation for ɛ- and ɔ-stem nouns and non-i-stem transitives, whose final vowel always seems to be reconcilable with the POC final vowel.

Two examples of the reduplication of a synchronic VCV show what I believe to be a reflex of a POC final vowel:

| | | | | |
|---|---|---|---|---|
| *tunu | *cook* | > | iniin | *cook* |
| *tumpu(?) | *grandparent..* | > | ipwiipw | *parentage* |

Thus:

$V_1C_1V_2$
$V_1C_1V_2V_1C_1V_2$ (reduplication)
$V_1C_1V_2V_1C_1$ (final vowel deletion)

In the two examples cited, $V_1 = V_2$. More evidence of this phenomenon is needed.

As with final vowels, it appears that all POC final consonants were lost in MOK. Of the nine POC consonants reconstructed in final position (*m, n, ŋ, p, t, k, ?, s, R), six (*n, p, t, ?, s, R) are regularly or often reflected as zero. Zero reflexes for *m, *ŋ, and *k are attested only in final position:

| | | | | |
|---|---|---|---|---|
| *tan(ou)m | *bury* | > | jɔwjɔw | *grave* |
| *awaŋ | *mouth* | > | aw | *mouth* |
| *manaŋ | *power* | > | manman | *possessing spiritual power* |
| *pakiwak | *shark* | > | pako | *shark* |
| *putik | *pull out* | > | wij | *pull out* |

A rule of final consonant deletion is supported by the existence of transitive-intransitive and adjective-noun doublets, where the deleted consonant appears in an environment where it would have been protected by a vowel suffix (later deleted). Thus:

| | | | | |
|---|---|---|---|---|
| *paŋ(ou)n | *awake, arouse* | > | pɔŋpɔŋ]$_{int}$ | *call* |
| | | | paŋin]$_{tr}$ | |
| *mpulu(t) | *gum,...* | > | pwil | *gum* |
| | | | pwilij | *to paste* |
| *putik | *pull out* | > | wij | *pull out* |
| | | | ? wijik | *carry* |

Note also that these examples suggest that consonant deletion preceded vowel deletion (whether diachronically or synchronically is open to question).

| | | |
|---|---|---|
| **paŋinV | paŋin | |
| | paŋi | (final C deletion) |
| paŋin | paŋ | (final V deletion) |

(Vowel raising and reduplication have been ignored here.)

It is interesting that the noun paradigms do not generally reflect the lost consonants. Whether this is to be taken as an instance of restructuring (analogical levelling?) or as an indication that the morphological process of possessive suffixation is 'newer' than the transitive-intransitive distinction is still quite speculative.

Final consonants are reflected in non-alternating forms in the following cases:

| | | | | |
|---|---|---|---|---|
| *inum | *drink* | > | nim | *drink* |
| *mpu(dr)i(t) | *excrement* | > | pwirej | *dirt* |
| *?uda(ŋ) | *lobster* | > | wuronna | *lobster* |

I must assume that the first two cases reflect a deleted suffix. The third seems to be cognate, but the final na remains unidentified. The geminate may have resulted from an *ŋ that has assimilated to the point of articulation of a following n.

## 3. CONSONANTS

The regular reflexes of POC consonants are summarised below. Cases of loss are not treated unless they appear to be regular reflexes in some non-final environment.

(All environments are tentative.)

| | | |
|---|---|---|
| *p | → | p/___a |
| | → | Ø |
| *mp | → | pw |
| *ŋp | → | pw |
| *t | → | Ø/___V |
| | | [+hi, -lo] |
| | → | j |
| *nt | → | s |
| *d | → | r |
| *nd | → | s |
| *k | → | k |
| *ŋk | → | k |
| *? | → | Ø |
| *m | → | m |
| | | mw/___u |
| *ŋm | → | mw |
| *n | → | n |
| *ŋ | → | ŋ |
| *l | → | l |
| *s | → | t |
| *ns | → | t |
| *R | → | Ø |
| *w | → | w |

### 3.1. Nasal grade

Reconstructions of POC non-sonorants (*p, d, t, s, k) are given in both oral and nasal grades on the assumption that these are reflected differently in the daughter languages. While this seems justified for *p and *mp, *d and *nd, and *t and *nt, my data give me no reason to believe this to be so for *s and *ns and *k and *ŋk. The data do suggest, however, that in a few instances MOK reflects the opposite grade from that given in the reconstructions. An incomplete list of these cases is given here.

reflex of *mp for *p

| | | | | |
|---|---|---|---|---|
| *puku | *knot* | > | pwuk | *knot* |
| *upu | *small plant* | > | upw | *unripe coconut* |

reflex of *p for *mp

| | | | | |
|---|---|---|---|---|
| *mpampa | *board* | > | pɔp | *board* |
| *tampi | *bowl* | > | japi | *bowl* |

reflex of *nd for *d

| | | | | |
|---|---|---|---|---|
| *daRa(?) | *blood* | > | nsa$^{4}$ | *blood* |
| *(dl)iki | *small* | > | siksik | *small* |

### 3.2. Laryngeals

*? > Ø

| | | | | |
|---|---|---|---|---|
| *?ate | *liver* | > | ɔj | *liver* |
| *?atop | *thatch* | > | ɔj | *thatch* |
| *?umu | *oven* | > | umw | *oven* |
| *da?a(n) | *branch* | > | ra | *branch* |
| *ma?udi(p) | *life,...* | > | mour | *life* |
| *pitu?u | *star* | > | uju | *star* |
| *ta?u | *year,...* | > | joonpar | *year* |

However, in the following examples *? seems to be reflected as w. Glide formation will be discussed more fully in Section 5).

| | | | | |
|---|---|---|---|---|
| *?una | *fish scale* | > | win | *hair, feathers, scales* |
| *?usa | *rain* | > | wut | *rain* |
| *?uluŋa | *pillow* | > | wiliŋ | *rest one's head* |
| *?uda(ŋ) | *lobster* | > | wuronna | *lobster* |
| *?uta(n) | *inland* | > | ɛwij | *inland* |

### 3.3. Velars

*k > k

| | | | | |
|---|---|---|---|---|
| *kuli(t) | *skin,...* | > | kil | *skin* |
| *kau | *tree* | > | suukɔ | *tree* |
| *kuku | *nail,...* | > | kik | *nail* |
| *kopu(t) | *mist,...* | > | sɔɔkoy | *fog* |
| *kuRita | *octopus* | > | kiij | *octopus* |

| | | | | |
|---|---|---|---|---|
| *kima | *giant clam* | > | kim | *kind of clam* |
| *ka(dr)u | *scratch* | > | karpwit | *scratch* |
| *iku | *tail* | > | ik | *tail* |
| *pinsiko | *flesh* | > | utuk | *flesh* |
| *tuki | *pound* | > | juk | *pound* |
| *nsakaRu | *reef* | > | tɔkɔ | *island* |
| *loku | *bend* | > | loklok | *bent* |
| *laki | *great* | > | lɛklɛkin | *big* |
| *pakiwak | *shark* | > | pako | *shark* |
| *mataku(t) | *afraid* | > | mijik | *fear* |

More problematic cognates are:

| | | | | |
|---|---|---|---|---|
| *ka(Rdr)aŋpa | *new* | > | kapw | *new* |
| *kau | *fish-hook* | > | kaw | *fosh-hook* |

In the first case, *aRa should be reflected as aa. The second suggests a reconstruction like *kawa.

*ŋk > k

| | | | | |
|---|---|---|---|---|
| *ŋkala | *male genitals* | > | kɔlɔ | *his penis* |
| *ŋkinit | *pinch* | > | kini | *pinch* |
| *saŋka(?) | *hop,...* | > | tak]$_{int}$ takɔ]$_{tr}$ | *drive a vehicle* |

## 3.4. Palatals and apicals

*t > j

| | | | | |
|---|---|---|---|---|
| *tama | *father* | > | jamaa | *his father* |
| *ta(n)si | *sea,...* | > | jɛt | *sea, salt* |
| *tolu | *three* | > | jiliw | *three* |
| *tuki | *pound* | > | juk | *pound* |
| *tapu- | *conch* | > | jowi | *conch* |
| *ta?u | *year* | > | joonpar | *year* |
| *taku | *back* | > | likinjarki- | *back* (cf. TRK sèkůr) |
| *taŋi(s) | *weep* | > | jɔŋ | *weep* |
| *?ate | *liver* | > | ɔj | *liver* |
| *mata | *eye* | > | maj | *eye* |
| *pitu | *seven* | > | ijiw | *seven* |
| *pitu?u | *star* | > | uju | *star* |
| *ŋmata | *worm* | > | mwaj | *worm* |
| *matolu | *thick* | > | mojul | *thick* |
| *mataku(t) | *afraid* | > | mijik | *fear* |

Note also *mputo *navel* > pwujjɔ *his navel*, showing gemination, and *kutu *louse* > pakit *delouse*, showing the irregular reflex *t > t.

*t > Ø

Both *t- and *-t- are often reflected as Ø when followed by a high vowel (ambiguously synchronic or diachronic).

| | | | | |
|---|---|---|---|---|
| *tina | *mother* | > | inaa | *his mother* |
| *tumpu(?) | *grandparent* | > | ipwiipw | *parentage* |
| *tu?u(d) | *stand up* | > | uuta | *stand up* |
| *tuRi(a) | *to string* | > | ir | *string* |
| *tunu | *cook* | > | iniin | *cook* |
| *tido | *look at* | > | iroŋ | *look at* |
| *anitu | *ghost* | > | eni | *spirit* |
| *matudu(R) | *sleep* | > | mɔir | *sleep* |
| *natu | *child* | > | naa | *his small thing* |

In one case *t > Ø before *e.

| | | | | |
|---|---|---|---|---|
| *mate | *die* | > | mee-ti | *die (animates)* |
| | | > | mɛɛ-ta | *die (plants)* |

If both these forms are taken to be cognate, the divergent development of *a must be explained in terms of the synchronic environment; that is, the suffixes -ti and -ta. It is tempting to try to extend the environment for a zero reflex of *t to non-low vowels in light of this cognate. However, in the examples above *t does not go to zero before *o. An explanation in terms of the reflexes of *e is hindered by the paucity of cognates reflecting that segment.

Exceptions to the high vowel environment are the following:

| | | | | |
|---|---|---|---|---|
| *pitu | *seven* | > | ijiw | *seven* |
| *pitu?u | *star* | > | uju | *star* |
| *tuki | *pound* | > | juk | *pound* |

A [j] reflex of *t is also found before a synchronic high vowel in:

| | | | | |
|---|---|---|---|---|
| *tolu | *three* | > | jiliw | *three* |
| *matolu | *thick* | > | mojul | *thick* |
| *mataku(t) | *afraid* | > | mijik | *fear* |

Rather than increase the list of exceptions, I propose to restrict the environment for a zero reflex of *t to $/\_\_\_\begin{bmatrix}*v\\+hi\end{bmatrix}$. For the first three exceptions I can offer no explanation.

*nt > s

| | | | | |
|---|---|---|---|---|
| *pu(n)ti | *banana* | > | wus | *banana* |
| *(dr)on(st)o | *night* | > | ros | *darkness* |
| *-nta | *our*, incl | > | -sa | *our*, dual incl |

*nt > Ø

| | | | | |
|---|---|---|---|---|
| kianto | *outrigger boom* | > | kia | *outrigger boom* |

Since the Mokilese maintain that their canoe terms are borrowed from Marshallese (MRS) (viz. MRS kiyęy *outrigger boom*), we can write this item off as a MRS loan. However, if other languages show similar reflexes for this item or if loss of *nt is not regular in this environment in MRS, then it must remain a problem.

*d > r

| | | | | |
|---|---|---|---|---|
| *doŋo | *hear* | > | roŋ | *hear* |
| *dua | *two* | > | riɔw | *two* |
| *da?a(n) | *branch* | > | ra | *branch* |
| *deŋa | *yellow* | > | rɔŋrɔŋ | *yellow* |
| *madama | *moon* | > | maram | *light* |
| *-da | *their* | > | -ra | *their*, dual |
| *matudu(R) | *sleep* | > | mɔyr | *sleep* |
| *pada | *pandanus* | > | kipar | *pandanus* |

The ambiguous reconstruction *(dr) is also reflected as r:

| | | | | |
|---|---|---|---|---|
| *(dr)an(i) | *day* | > | rɛɛn | *day* |
| *(dr)on(st)o | *night* | > | ros | *darkness* |
| *ka(dr)u | *scratch* | > | karpwit | *scratch* |

*nd > s

| | | | | |
|---|---|---|---|---|
| *ndau(n) | *leaf* | > | sɔ | *leaf* |
| *ndaŋma | *forehead* | > | sɔmwɛ | *his forehead* |
| *ndamu | *chew betel,...>* | | samwɛ]$_{tr}$ | *lick* |

*s > t

| | | | | |
|---|---|---|---|---|
| *solo | *mountain* | > | tol | *mountain* |
| *sili | *enter* | > | til | *enter* |
| *sulu | *torch* | > | til | *torch* |
| *saŋka(?) | *hop,...* | > | tak]$_{int}$ | *drive a vehicle* |
| *?usa | *rain* | > | wut | *rain* |
| *masawa | *sea* | > | mataw | *open sea* |
| *masaki(t) | *pain* | > | mɔtɔk | *pain* |
| *ta(n)si | *sea,...* | > | jɛt | *sea, salt* |

*s > Ø

| | | | | |
|---|---|---|---|---|
| *sala(n) | *road* | > | al | *road* |
| *tasimi | *sharpen* | > | jaim | *sharpen* |
| *suli | *shoot* | > | il | *small plant* |
| *sida | *they* | > | iir | *they* |

*ns > t

| | | | | |
|---|---|---|---|---|
| *nsama | *outrigger* | > | tam | *outrigger* |
| *nsiwa | *nine* | > | tiwɔw | *nine* |
| *ansa(n) | *name* | > | at | *name* |
| *nsakaRu | *reef* | > | tɔkɔ | *island* |
| *nsoka | *stab* | > | tok]$_{int}$ | *stab* |
| | | > | tɔkɔ]$_{tr}$ | *stab* |
| *mansu(rR) | *plenty* | > | mɔt | *fat* |
| *pinsiko | *flesh* | > | uduk | *flesh* |

*ns > Ø

| | | | | |
|---|---|---|---|---|
| *nsaŋi | *wind* | > | ɛŋ | *wind* |

The environments for a zero reflex of both *s and *ns overlap with those of the reflex t:

| *(n)s > t | | *(n)s > Ø | |
|---|---|---|---|
| *saŋka(?) | *hop,...* | *sala | *road* |
| *sulu | *torch* | *suli | *shoot* |
| *ta(n)si | *sea,...* | *tasimi | *sharpen* |
| *nsakaRu | *reef* | *nsaŋi | *wind* |

Since *s and *ns do not appear to be reflected differently it is possible to assign one reflex to each of them and explain the overlap as a change of 'grade'. Given the preponderance of t reflexes for both *s and *ns (eight out of twelve for *s and nine out of ten for *ns), however, this solution seems quite ad hoc. If, however, all the MC languages were to agree as to which items reflecting *s or *ns share a given reflex, a grade-alternation solution could more realistically be proposed for Proto-Micronesian (PMC).

*n > n

| | | | | |
|---|---|---|---|---|
| *niu | *coconut* | > | ni | *coconut* |
| *napo | *surf* | > | no | *wave* |
| *nua | *inside* | > | nɛɛn | *inside* |
| *anitu | *spirit* | > | eni | *ghost, spirit* |
| *inum | *drink* | > | nim | *drink* |
| *-na | *his* | > | -n | *his* |
| *onom | *six* | > | wonow | *six* |
| *ŋkinit | *pinch* | > | kini | *pinch* |
| *manaŋ | *power* | > | manman | *spiritually powerful* |

*n > Ø

| | | | | |
|---|---|---|---|---|
| *namu(k) | *mosquito* | > | amwijɛ | *mosquito* |
| *pan(i) | *wing* | > | pɔ | *arm, wing* |
| *ponu | *turtle* | > | woy | *turtle* |
| *tanum | *bury* | > | jɔwjɔw | *grave* |

With the exception of the first example of loss of *n, this change seems to have taken place in the environment */V__V#, , where * indicates a
[+ba][+hi]
proto-language environment, and where the condition on the preceding vowel is included because of

| | | | | |
|---|---|---|---|---|
| *ŋkinit | *pinch* | > | kini | *pinch* |
| *inum | *drink* | > | nim | *drink* |

and the condition # is included because of

| | | | | |
|---|---|---|---|---|
| *anitu | *spirit* | > | eni | *spirit, ghost* |

This environment is nothing more than an observation; the phonetic motivation of such a change is not clear to me.

*l > l

| | | | | |
|---|---|---|---|---|
| *laŋi | *sky* | > | lɔŋ | *sky* |
| *laŋo | *fly* (n.) | > | lɔŋ | *fly* (n.) |
| *lima | *five* | > | limɔw | *five* |
| *loku | *bend* | > | loklok | *bent* |
| *mpulu(t) | *gum,...* | > | pwil | *gum, sap* |
| *tolu | *three* | > | jiliw | *three* |
| *alap | *take* | > | ɔlɔ]$_{tr}$ | *take* |

I have also found some evidence for a zero reflex of *l in the environment */ V [-lo] :

| | | | | |
|---|---|---|---|---|
| *malu | *soft* | > | mo | *grass* |
| *maliŋ | *sour* | > | mɛɛn | *sour* |
| *ŋmalo | *submerge* | > | mwoiti | *sink* |

This is somewhat tempting since the environment postulated is reminiscent of that for *n > Ø and, to a lesser extent, that of *t > Ø. (Note that the feature [-lo], rather than [+hi], appears here to account for *ŋmalo. A similar solution was considered for the zero reflex of *t at the beginning of this section.) The first, while regular phonologically, is semantically odd. The long vowel and final n of the second case and the front glide in mwoi- in the third would require further explanation. In any event, waliw *eight* < *walu *eight* and jɔl *rope* < *tali *rope* would remain exceptional.

*R > Ø[5]

| | | | | |
|---|---|---|---|---|
| *ka(Rdr)aŋpa | *new* | > | kapw | *new* |
| *RuŋRa | *house* | > | umw | *house* |
| *daRa(?) | *blood* | > | nsa | *blood* |
| *nsakaRu | *reef* | > | tɔkɔ | *island* |
| *kuRita | *octopus* | > | kiij | *octopus* |
| *paRi | *rayfish* | > | pɔ | *ray* |
| *Ripa | *near* | > | ipaa | *near him* |

*R > r

| | | | | |
|---|---|---|---|---|
| *?apaRa | *shoulder* | > | aparɔ | *his shoulder* |
| *tuRi(a) | *to string* | > | ir | *string* |

Grace has proposed (in a 1972 seminar) that the zero reflex of *R is conditioned by a following high vowel. This seems to be a reasonable hypothesis, given the similarity between this environment and those proposed for a zero reflex of the other protoapicals. The problem with this solution in light of the MOK evidence is a statistical one. Of the nine cognates showing reflexes of *R, only two have the r reflex and one of these shows it in the environment proposed for the zero reflex. Thus I cannot present any evidence from MOK that strongly supports this treatment of *R.

## 3.5. Labials

*p > p

| | | | | |
|---|---|---|---|---|
| *pan(i) | *wing* | > | pɔ | *arm, wing* |
| *pakiwak | *shark* | > | pako | *pako* |
| *pani | *bait* | > | pɛɛn | *bait* |
| *pili(?) | *select* | > | pil | *choose* |
| *pi(dr)i | *twist* | > | kapir | *twist* |
| *manipi(s) | *thin* | > | minipnip | *thin* |
| *Ripa | *near* | > | ipaa | *near him* |

*p > Ø

| | | | | |
|---|---|---|---|---|
| *pinsiko | *flesh* | > | utuk | *flesh* |
| *pitu?u | *star* | > | uju | *star* |
| *puko | *net* | > | uk | *net* |
| *punti | *banana* | > | wus | *banana* |
| *ponu | *turtle* | > | woy | *turtle* |
| *puki | *return* | > | wik | *change* |
| *api | *fire* | > | ɔy | *fire* |
| *napo | *surf* | > | no | *wave* |
| *mapo | *heal* | > | mo | *heal* |
| *kopu(t) | *dust, mist* | > | sɔɔkoi | *fog* |
| *tapu(n)i | *bury* | > | jɔwjɔw | *grave* |
| *tapu- | *conch* | > | jowi | *conch* |

(Apparent cases of *p > w will be treated under Glide Formation in Section 5.)

The normal reflex of *p seems to be zero, except before *a. Of the eleven p-reflexes of *p, three show retention before *i: *pili? *select* > pil *choose*, *pi(dr)i *twist* > kapir *twist*, and *manipi(s) *thin* > minipnip *thin*, as compared with *api *fire* > ɔy *fire* and *pitu *seven* > ijiw *seven*. I have no explanation for these cases.

In 3.1. above I treat the p of japi *bowl* as an irregular-grade reflex of *tampi *bowl*. It is evident, however, that p is also an exceptional reflex of *p in this environment. An alternative explanation for this form is given below.

*p > pw

| | | | | |
|---|---|---|---|---|
| *puku | *knot* | > | pwuk | *knot* |
| *upu | *young leaves* | > | upw | *unripe coconut* |

In 3.1. these are treated as examples of an *mp reflex for *p. An alternate hypothesis might have had *p > pw/___*u $\{{}^{\#}_{[+back]}\}$. This rule would have had to apply before final vowel deletion, since a final *u provides part of its environment. Possible exceptions are:

| | | | | |
|---|---|---|---|---|
| *tapu(n)i | *afterbirth* | > | jɔu | *afterbirth* |
| *kopu(t) | *mist,...* | > | sɔɔkoi | *fog* |
| *puki | *return* | > | wik | *turn, change* |
| *puko | *net* | > | uk | *net* |

The first three might be explained by a more detailed specification of 'velar (rounded?) environment', whereby the $[{}^{C}_{+back}]$ must be followed by *u at the time of application of the rule. While there appears to be no synchronic kw in MOK, it is not inconceivable that a historical *ku could have left synchronic reflexes of its presence in the form of velarised labials and/or rounded vowels. The potential of this hypothesis as an explanation of such reflexes is yet to be explored. I shall return briefly to this 'velar environment hypothesis' in the discussion of high vowels.

*m > m

> mw */___u

In my data *m is found before only two vowels, *a and *u. In the former environment it is reflected as m, in the latter as mw.

| | | | | |
|---|---|---|---|---|
| *madama | *moon* | > | maram | *light* |
| *mata | *eye* | > | maj | *eye* |
| *mapo | *heal* | > | mɔ | *heal* |
| *tama | *father* | > | jamaa | *his father* |
| *lima | *five* | > | limɔw | *five* |
| *nsama | *outrigger* | > | tam | *outrigger* |
| *muta(?) | *vomit* | > | mwwuj | *vomit* |
| *mudi | *behind* | > | mwirin | *after* |
| *ʔumu | *oven* | > | umw | *oven* |
| *namu(k) | *mosquito* | > | amwijɛ | *mosquito* |
| *limu | *seaweed* | > | limw | *moss, seaweed* |
| *ndamu | *chew betel* | > | samwɛ | *lick* |

*mp > pw

| | | | | |
|---|---|---|---|---|
| *mpo- | *smell* | > | pwoɔ | *his smell* |
| *mpua | *areca* | > | pwu | *betel nut* |
| *mpulu(t) | *gum..* | > | pwil | *gum, sap* |
| *mputo | *navel* | > | pwujjɔ | *his navel* |
| *tumpu(?) | *grandparent* | > | ipwiipw | *parentage* |
| *mpu(dr)i(t) | *excrement* | > | pwirej | *dirt* |
| *mponot | *blocked* | > | pwon | *blocked, stuffed* |
| *mpou | *post* | > | pwo | *pole* |

In all these examples, a pw reflex of *mp is found in the environment */___$\begin{bmatrix} V \\ +back \\ +rnd \end{bmatrix}$. In 3.1. p reflexes of *mp were treated as grade alternants. It is perhaps worth pointing out that the three examples in question show *mp/___$\begin{bmatrix} V \\ -rnd \end{bmatrix}$ which suggests that these discrepancies may result more from other environmental factors than from grade alternation.

## 3.6. Labio-velars

*ŋp > pw

| | | | | |
|---|---|---|---|---|
| *ŋpoŋi | *night* | > | pwoŋ | *night* |
| *ka(Rdr)aŋpa | *new* | > | kapw | *new* |

*ŋm > mw

| | | | | |
|---|---|---|---|---|
| *Ruŋma(?) | *house* | > | umw | *house* |
| *ŋmata | *worm* | > | mwaj | *worm* |
| *ndaŋma | *forehead* | > | sɔmwɔ | *his forehead* |

*w > w

| | | | | |
|---|---|---|---|---|
| *awaŋ | *mouth* | > | aw | *mouth* |
| *masawa | *ocean* | > | mataw | *ocean* |
| *wawɔ | *forest* | > | wɔwɔ | *scrub* |

### 3.7. Some Generalisations Concerning Consonant Loss in Nonfinal Environments

The following consonants show loss as a regular reflex: *p, t, *n, s, l(?), R, ?. The nasal grade of *s and *t each show one cognate with a zero reflex. With the exception of *? (and perhaps *R), all these consonants are [+ant]. *? is aberrant in that its only reflex is zero. Of the anterior consonants in the POC inventory only *m, *d (*r), and *nd do not have zero as a reflex. If *d (*r) had already become a retroflex before POC anterior consonants began to be lost, this would explain the fact that it did not participate in this change. As pointed out earlier, *m is found only in very restricted environments, which might explain the apparent lack of zero reflexes. As regards *R, I have no reliable information on its probable point of articulation.

In most cases the environments for loss are quite similar, suggesting a $[^{*V}_{+hi}]$ as the prime conditioning factor. One should bear in mind that in none of these cases are the environments exceptionless and in most, if not all, the phonetic motivation is unclear. ((*) indicates an ambiguous synchronic POC environment.)

*p > Ø / ___ $[^{V}_{-lo}]$

*R > Ø / ___ $[^{V}_{+hi}]$ (from Grace; see Section 3.4)

*t > Ø / ___ $[^{(*)\ V}_{+hi}]$

*s > Ø / sporadically

n > Ø / $[^{(*)\ V}_{+back}]$ ___ $[^{V}_{+hi}]$

l > Ø / a___ $[^{V}_{-lo}]$

These observations are intended to be suggestive only. In my opinion no real generalisations can be made about these cases of loss (nor, for that matter, about nasal grade variation) until a comparison has been made of such phenomena throughout MC. Only when it has been determined which cognate items and segments reflect similar changes in the same environments can we begin to draw any conclusions about what constitutes a regular development in a particular MC language.

## 4. VOWELS

It is trivial to point out that historical investigation cannot be carried on in a vaccuum. A comprehensive account of any diachronic phenomenon must be built upon a thorough analysis of the synchronic facts. The sketchy nature of the following treatment of MOK vowels reflects an as yet merely superficial understanding of patterns of vowel alternation as they are evidenced in the possessive paradigm and in various aspects of the verb morphology. I will draw upon such observations as I have been able to make, though it must be kept in mind that any conclusions must be considered tentative.

## 4.1. *a

*a > a

### 4.1.1. * /___Ca

| | | | | |
|---|---|---|---|---|
| *da?a(n) | *branch* | > | ra | *branch* |
| *mada | *fermented* | > | mar | *preserved breadfruit* |
| *tama | *father* | > | jamaa | *his father* |
| *mata | *eye* | > | maj | *eye* |
| *masawa | *sea* | > | mataw | *ocean* |

In other forms in the paradigms of inalienably possessed nouns like jamaa *his father* and maj *eye*, *a is reflected differently: jɔmɔy *my father*, jɛmɛn *father of*, and mijɔ *his eye*.

### 4.1.2. * / some open monosyllables

| | | | | |
|---|---|---|---|---|
| *-da | *their* | > | -ra | *their, dual* |
| *-nta | *our, incl* | > | -sa | *our, dual incl* |
| *-na | *his* | > | -n | *his* |

The vowel is preserved in the first two cases undoubtedly because at some stage in the history of MOK these suffixes were phonetically (morphologically) more complex. Note the plural forms -rai *their* and -sai *our incl*. If the vowel in the deleted syllable were inferred to be one that conditioned an a reflex of a preceeding *a, then these cases need not be treated differently from other a reflexes.

It is also possible that the above environmental statement might be generalised to all open monosyllables (including those resulting from final consonant deletion). Evidence of this is based solely on the form *pat *four* > paaw *four*, and this example is itself prejudiced by the existence of a second stem for *four*, as in ɔpɔŋ *(four (serial)*. It is not at all likely that a more thorough search might produce real counterexamples to this claim.

### 4.1.3. * /___mu

| | | | | |
|---|---|---|---|---|
| *namu(k) | *mosquito* | > | amwijɛ | *mosquito* |
| *ndamu | *chew betel* | > | samwsamw]$_{int}$ | *lick* |
| | | | samwɛ]$_{tr}$ | |

It might be that *u# was dropped when it velarised *m, so that these items were not vowel-final when raising applied. This is suggested by the fact that the intransitive of samwɛ (samwsamw) does not reflect a final vowel, but this argument carries no weight if reduplication occurred after final vowel deletion. In the case of amwijɛ *mosquito* the i is probably excrescent. This question will be touched upon briefly in the discussion of high vowels below.

### 4.1.4. Other isolated occurrences of * /___Cu#

| | | | | |
|---|---|---|---|---|
| *ka(dr)u | *scratch* | > | karpwit | *scratch* |
| *taku | *back* | > | likinjarki- | *back* (cf. TRK sèkůr *back*) |
| *walu | *eight* | > | waliw | *eight* |
| *manu(k) | *animal* | > | maan | *animal* |

I hesitate to draw any conclusions about these cases given examples like: *mataku *afraid* > mijik *fear*.

### 4.1.5. Other cases

| | | | | |
|---|---|---|---|---|
| *pakiwak | *shark* | > | pako | *shark* |
| *kianto | *outrigger boom* | > | kia | *outrigger boom* |
| *tampi | *bowl* | > | japi | *bowl* |

In spite of the problematic nature of several of the examples — metathesis in the case of -jarki- *back*, vowel length in maan *animal* (the vowels in monosyllabic nouns do not regularly lengthen in MOK, unlike those of Ponapean (PNP), TRK, etc.), and the grade alternation problems discussed above (Section 3) — the facts of Section 4.1. lead me to believe that an environment both velarised and rounded did contribute to a reflexes of *a in the examples cited. Firm evidence is still lacking however.

### 4.1.6. *a > ɛ

This reflex is attested in:

| | | | | |
|---|---|---|---|---|
| *tansi | *sea,...* | > | jɛt | *sea..* |
| *laki | *great,...* | > | lɛklɛkin | *big* |
| *nsaŋi | *wind* | > | ɛŋ | *wind* |
| *manu(k) | *animal* | > | mɛn | *cl. for animates* (cf. maan) |
| *(dr)ani | *day* | > | rɛɛn | *day* |
| *mate | *die* | > | mɛɛta | *die (plants)* |
| *ma(R)a | *ashamed* | > | mɛɛk | *ashamed* |
| *ndaŋma | *forehead* | > | sɔmwɛ | *his forehead* |
| *ʔuna | *scale* | > | winɛ | *its scales, feathers* |

The last two cognates probably reflect a deleted vowel suffix. The ɛ outcome suggests an earlier *a+i, but I am unable to give a more detailed justification of this conclusion here.

If mɛɛk *ashamed* is indeed cognate, then the -ɛk might be the 'middle' suffix attested synchronically. Nothing in the synchronic morphophonemics of -ɛk, or of any other suffix of this type, would lead one to expect mɛɛk < *ma + ɛk.

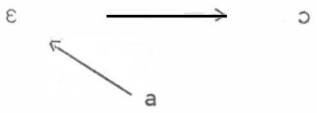
The final *i of four of the examples is a logical environment for *a > ɛ, but still leaves a residue of problem cases. Thus:

| | | | | |
|---|---|---|---|---|
| *taŋi(s) | *weep* | > | jɔŋ | *weep* |
| *pan(i) | *wing* | > | pɔ | *arm, wing* |
| *api | *fire* | > | ɔy | *fire* |
| *masaki(t) | *pain* | > | mɔtɔk | *pain* |

The case of rɛɛn *day* is further complicated by the irregular vowel length, and that of mɛn *animate classifier* by the existence of another reflex: maan *animal*.

While I can draw no conclusions from these data, it seems to me that a front environment in terms of consonants, perhaps more than vowels, was a contributing factor to ɛ reflexes of *a. In the case of ɛŋ *wind* and mɛɛta *die*, a consonant that normally yields 'front' reflexes was lost. (Comparing *nsaŋi > ɛŋ with *taŋi(s) > jɔŋ, one might get the impression that such loss constituted more of a front environment than did the presence of a fronting consonant.)

In comparing MOK with other nuclear MC languages one is impressed by the fact that ɔ seems to have proliferated in MOK at the expense of ɛ. Thus: MOK ɔy *fire*, PNP ɛɛy TRK ååf KSR eɛ, and MOK jɔmɔy *my father* PNP sɛmɛy MRS jema-. These facts, as well as the fact that the i reflexes of *a can best be explained by positing an intermediate ɛ-stage, lead one to conclude that the replacement of ɛ by ɔ is an innovation in MOK,




and that those items which retain ɛ were in some way excluded from the ɛ > ɔ change. That these conclusions are probably false will be shown in Section 4.1.9.

## 4.1.7. *a > e

This reflex is attested in only two cognates:

| | | | | |
|---|---|---|---|---|
| *anitu | *ghost* | > | eni | *ghost* |
| *mate | *die* | > | meeti | *die (for animates)* (cf. mɛɛta *die (plants)*) |

In view of the alternation in the two reflexes of *mate *die*, and similar forms (e.g. ŋɛn, ŋenin *spirit, spirit of*) and in light of the PNP reflex ɛni *ghost* < *anitu, I conclude that this reflex of *a is probably conditioned synchronically in MOK by a 'front' environment (in this case, an anterior consonant followed by i). My synchronic analysis of MOK is not yet sufficiently refined for me to be able to give a clearer statement.

## 4.1.8. *a > i

| | | | | |
|---|---|---|---|---|
| *mataku | *fear* | > | mijik | *afraid* |
| *mata | *eye* | > | mijɔ | *his eye* (maj *eye*) |
| *kami | *we, excl* | > | kimi | *we, excl* |
| *kamu | *you, pl* | > | kimwi | *you, pl* |
| *manipi(s) | *thin* | > | minipnip | *thin* |

In comparing the first two examples with the MRS and PNP cognates; mijak, meja- and mɛsɛk, mɛsɛ respectively, we find what I take to be an intermediate stage in the development of the MOK forms. It seems that where ɛ preceeded j it was raised to i in MOK in open syllables. The following k of

mijik must have provided enough of a 'high' environment to permit 'assimilatory' raising of the following ɛ as well. I would like to attribute the i reflexes in kimi and kimwi to a similar consonantal environment. The raising in the last cognate is anomalous in this analysis, however, and must be attributed to the influence of the two following i's.

### 4.1.9. *a > ɔ

This is by far the most common reflex of *a. It is found:

a) alternating with a in paradigmatically-related forms

| | | | | | | |
|---|---|---|---|---|---|---|
| *ansa(n) | *name* | > | at | *name* | ɔtɔ | *his name* |
| *alap | *take* | > | aljapw | *take land* | ɔlɔ | *take* |
| *awaŋ | *mouth* | > | aw | *mouth* | ɔwɔ | *his mouth* |

b) before *(C)V [-lo]

| | | | | |
|---|---|---|---|---|
| *ndau | *leaf* | > | sɔ | *leaf* |
| *matudu | *sleep* | > | mɔyr | *sleep* |
| *mansu(rR) | *plenty* | > | mɔt | *fat* |
| *api | *fire* | > | ɔy | *fire* |
| *laŋi(t) | *sky* | > | lɔŋ | *sky* |
| *pan(i) | *wing* | > | pɔ | *arm, wing* |
| *masaki(t) | *pain* | > | mɔtɔk | *pain* |
| *paRi | *ray* | > | pɔ | *ray* |
| *taŋi(s) | *weep* | > | jɔŋ | *weep* |
| *aŋoaŋo | *yellow* | > | ɔŋ | *yellow, turmeric bush* |
| *ʔatop | *thatch* | > | ɔj | *thatch* |
| *laŋo | *fly* | > | lɔŋ | *fly (n.)* |

c) other examples include

| | | | | |
|---|---|---|---|---|
| *mpampa | *board* | > | pɔp | *board* |
| *ŋkala | *male genitals* | > | kɔlɔ | *his penis* |
| *ndaŋma | *forehead* | > | sɔmwɛ | *his forehead* |

The last two examples are the third person singular of two inalienable nouns that are defective in that they do not possess isolate forms. If these are hypothesised to be **kal *penis* and **samw *forehead*, then these forms fit the same pattern as the examples under a) above. The first case under c) remains problematic.

A comparison of some MOK ɔ reflexes of *a with reflexes of the same items in PNP and MRS will show that the backing is in fact much more general. (It should of course be kept in mind that in MRS fronting/backing seems to be wholly dependent on the synchronic consonantal environment — see Bender 1968.)

| POC | PNP | MOK | MRS | |
|---|---|---|---|---|
| *api | ɛɛy | ɔy | — | *fire* |
| *laŋi(t) | laaŋ, lɛŋin | lɔŋ | lag | *sky* |
| *ansa(n) | aat, ɛtɛ | at, ɔtɔ | yat, yeta- | *name* |
| *ndau(n) | tɛɛ | sɔ | — | *leaf* |
| *masaki(t) | mɛtɛk | mɔtɔk | metak | *pain* |
| *matudu(rR) | mɛyr | mɔyr | majiẗ | *sleep* |

| | | | | |
|---|---|---|---|---|
| *natu | nɛi | nɔi | naji- | *child* |
| *pan(i) | pɛɛ | pɔ | pay | *arm, wing* |
| *tama | sɛmɛ- | jɔmɔ- | jema- | *father* |
| *taŋi(s) | sɛŋ | jɔŋ | jag | *cry* |
| *tina | inɛ- | inɔ- | jine- | *mother* |
| *yaŋo<br>aŋoɛŋo | ɔŋɔɔŋ | ɔŋ | — | *yellow, turmeric* |
| *ʔatop | ɔɔs | ɔj | haj | *thatch* |
| *ʔate | ɛɛ | ɔj | haj | *liver* |
| *laŋo | lɔɔŋ | lɔŋ | ʈagä | *fly (n.)* |
| nsakaRu | tɛkɛ | tɔkɔ | tekay | *reef* |

It should be pointed out that in MRS raising occurs only under very specific and restricted conditions, while in MOK and PNP raising accompanies all instances of fronting and backing, at least as these languages are being analysed at present. a-raising in MRS is a dissimilation rule of the following type:

aCa → eCa (Bender 1969)

A rule of this kind does not seem to be operating in MOK, at least not synchronically, since aCa sequences are permitted, as in takar *to illuminate*. Were it possible to construct a version of this rule that applied only to alternating forms, as in the possessive paradigms, the rule in question would have to be supplemented by a later rule of raising assimilation to account for the fact that in most cases every reflex of *a is raised in an item where raising has applied. A more refined version of this analysis probably could be produced after a closer examination of the data, but at present this solution does not seem the most likely of the alternatives.

A second possibility involves a rule of prejunctural raising followed by a leftward assimilation rule. Whether prejunctural raising should be ordered before or after final vowel deletion seems to rest on the theoretical admissibility of 'dangling junctures'. Other details of this analysis remain to be worked out.

A final alternative would have raising apply before a following non-low vowel and proceed (with some restrictions) leftward through the word. Synchronically this analysis has the disadvantage of forcing us to make a statement about the height of a vowel that never appears on the surface. In historical terms, however, it seems to be the correct analysis (as suggested in Section 4.1.9(b)).

I should like to make clear that I am treating the environment of raising and the phonetic outcome of raising as two different problems. I have no conclusions to draw about the former, other than those sketched above. (For convenience, a comparison of the three possible rule sequences follows.) The rest of this section will be devoted to a consideration of the problem of the relative backness of raised vowels.

I. Raising Dissimilation

| **ata | | **ata+V | | |
|---|---|---|---|---|
| at | | ata+ | | Final Vowel Deletion |
| — | | ɔta+ | | Raising Dissimilation |
| — | | ɔtɔ+ | | Raising Assimilation |
| at | *nɑme* | ɔtɔ | *his nɑme* | |

II. Prejunctural Raising

| | | | | |
|---|---|---|---|---|
| **ata | | **ata+V | | |
| ata | | atɔ+V | | Prejunctural Raising |
| at | | atɔ+ | | Final Vowel Deletion |
| — | | ɔtɔ+ | | Assimilation |
| at | *name* | ɔtɔ | *his name* | |

III. Vowel-sensitive Raising

| | | | | |
|---|---|---|---|---|
| **ata | | **ata+V<br>[-lo] | | |
| ata | | atɔ+V | | Raising |
| at | | atɔ+ | | Final Vowel Deletion |
| — | | ɔtɔ | | Assimilation |
| at | *name* | ɔtɔ | *his name* | |

As is evident from the chart given earlier in this section, the cognates in which MOK and PNP agree as to the backness of the reflex of *a are restricted to those cases in which *a was followed in the next syllable by *o.

In previous discussion it was assumed that raising as a historical rule followed fronting/backing. This need not have been the case, however. In the alternative solution, raising would apply before backness adjustment.

Backing First

ɛ ɔ
↑ ↑
ä ← a → ä̈

Raising First

ɛ ← ʌ → ɔ
↑
a

In formal terms, the raising-first solution is simpler in that at the time raising applies there is only one $\begin{bmatrix} V \\ -hi \\ +lo \end{bmatrix}$, while with the backing first solution there are three, necessitating the addition of the features $\begin{bmatrix} \alpha back \\ \alpha rnd \end{bmatrix}$ to the input of the rule to prevent a itself from raising. This solution is further complicated by the fact that it forces us to make the 'unnatural' assumption that a) in MOK, certain instances of *a backed even before front vowels (e.g. *api > ɔy), and that b) in PNP, certain instances of *a were fronted even before back vowels (e.g. *natu > nei). (The latter is less crucial since we know that *a became ɔ before *o in both languages, and that there is evidence that points to a fronting of *u).

In addition to these advantages, the raising-first solution provides the beginnings of an explanation for the lack of agreement between MOK and PNP on the position of the reflex of raised *a. If raising had been from [a] to [ʌ], then this disparity is explained by the fact that in MOK there was a fronting of [ʌ] in a 'front' environment (which will probably turn out to require the presence of both a front vowel and a high consonant), while in PNP [ʌ] was backed in a 'back' environment. In MOK, all those occurrences of [ʌ] that were not fronted were interpreted phonemically as ɔ (merging with lowered reflexes of *a), while in PNP those occurrences of [ʌ] that were not backed became phonemically ɛ. This analysis might turn out to be reconcilable with the backing-first solution set forth above, but I am unable to explore this and other refinements at this time.

### 4.1.10. *a > o

This reflex appears in the environment * /____ $\begin{Bmatrix} ? \\ p \end{Bmatrix}$ V [+rnd] , as in:

| | | | | |
|---|---|---|---|---|
| *ma?udi(p) | *life,...* | > | mour | *life* |
| *napo | *surf* | > | no | *wave* |
| *mapo | *heal* | > | mo | *heal* |
| *tapu- | *conch* | > | jowi | *conch* |
| *ta?u | *year,...* | > | joonpar | *year* |

I assume that the loss of *? and *p in this position resulted in the creation of a glide agreeing in backness and roundness with the following vowel If that vowel was final, it was later lost. The fact that in these examples the hypothesised glide seems to appear in only one form, jowi *conch* — it might be present, though difficult to hear, in mour *life* as well — suggests that under certain conditions these glides were deleted (see Section 5).

A more thorough analysis might shed light on the relation between the conditioning of this rule and that of the rules accounting for the other reflexes of *a, perhaps in terms of a rounded environment, as opposed to the palatal and front and back environments suggested above. Any such synthesis is still quite speculative however.

*a is also reflected as o in:

| | | | | |
|---|---|---|---|---|
| *matolu | *thick* | > | mojul | *thick* |
| *?uda(ŋ) | *lobster* | > | wuronna | *lobster* |

I have no explanation to offer for these reflexes.

## 4.2. *e

*e > $\begin{Bmatrix} ɔ \\ ɛ \end{Bmatrix}$

Only three cognates show reflexes of *e:

| | | | | |
|---|---|---|---|---|
| *deŋa | *tumeric,...* | > | rɔŋrɔŋ | *yellow* |
| *kie | *pandanus* | > | kiɔ | *his mat* |
| *mpeŋka | *bat* | > | pwɛɛk | *bat* |

The first two examples suggest that *e merged with the raised reflexes of *a. This type of solution might be possible for the third case as well, although the environment for fronting is not obvious and the long vowel is problematic. It is impossible to say more with such scant data.

## 4.3. *o

This vowel is most commonly reflected as o.

| | | | | |
|---|---|---|---|---|
| *ŋpoŋi | *night* | > | pwoŋ | *night* |
| *ponu | *turtle* | > | woy | *turtle* |
| *loku | *bent* | > | loklok | *bent* |
| *mpou | *post* | > | pwo | *pole* |

| | | | | |
|---|---|---|---|---|
| *kopu(t) | *mist* | > | sɔɔkoi | *fog* |
| *doŋo | *hear* | > | roŋ | *hear* |
| *(dr)on(st)o | *night* | > | ros | *darkness* |
| *mpo- | *smell* | > | pwoɔ | *his smell* |
| *solo | *mountain* | > | tol | *mountain* |
| *tido | *look at* | > | iroŋ | *look at* |
| *mponot | *blocked* | > | pwon | *blocked* |

In the above example, *o is either:

a) followed by a round vowel and/or

b) preceded by a consonant that yields a rounded reflex and/or

c) followed by a back consonant.

These are all environments in which our previous assumptions about both the synchronic and diachronic phonology of MOK would lead us to expect a round vowel. I might add that synchronically o seems to be the vowel most responsive to its environment in terms of backness. Thus, the 'o' in pwo *pole* is phonetically [o], while that in tol *mountain* is [ə], and that in woin, construct of woi *turtle*, begins back and round but through most of its articulation sounds much more like [e]. It is perhaps odd that the o of woi is not equally fronted, unless there is something about the construct state that provides more of a fronting environment (compare jɔmɔi *my father* and jɛmɛn *father of*). This would appear to coincide with the PMC reconstruction *-ni for the construct, but must be examined in more detail.

*onom *six* yields wonow *six, general classifier* and oon *six, general counting.* The latter is probably morphemically complex, consisting of the numeral base **on *six* and a prefix, tentatively **a, which is variously realised as ɔ (ɔpɔŋ *four*, a (alim *five*), e (eij *seven*), and in this example as o. I cannot yet explain the presence of the initial [w] in the unprefixed form.

One alternation suggests that a reflex of *o was lowered before a reflex of *a at some later stage following final vowel deletion:

| | | | | |
|---|---|---|---|---|
| *nsoka | *stab* | > | tok]$_{int}$ | *stab* |
| | | | tɔkɔ]$_{tr}$ | |

*o is reflected as a high vowel in two examples:

| | | | | |
|---|---|---|---|---|
| *tolu | *three* | > | jiliw | *three, general cl* |
| *matolu | *thick* | > | mojul | *thick* |

The environment for this raising seems to coincide with that for the rule *a > i (see Section 4.1.8.). An interesting problem, however, is that while it does not participate in the *a > i change, PNP does show raising in these forms: sili- *three* and mɔsul *thick*. The difference in backness of the high vowel reflexes in these two examples is also without an obvious explanation.

## 4.4. *i

*i > i

| | | | | |
|---|---|---|---|---|
| *niu | *coconut* | > | ni | *coconut tree* |
| *anitu | *ghost* | > | eni | *ghost* |
| *tina | *mother* | > | inaa | *his mother* |

| | | | | |
|---|---|---|---|---|
| *lima | *five* | > | limɔw | *five, general cl* |
| *ŋkinit | *pinch* | > | kini | *pinch* |
| *limu | *seaweed* | > | limw | *moss, seaweed* |
| *sili | *enter* | > | til | *enter* |
| *pitu | *seven* | > | eij | *seven, serial* |
| *tampi | *bowl* | > | japi | *bowl* |

*i > u / ___ V
[+rnd]

| | | | | |
|---|---|---|---|---|
| *pinsiko | *flesh* | > | utuk | *flesh* |
| *pitu?u | *star* | > | uju | *star* |

As is obvious from the examples that show i reflexes of *i many instances of *u do not cause backing of a preceding *i, contrary to what the above environmental statement would suggest. In this section counterexamples to this claim are considered and some general solutions suggested.

In the case of *limu *seaweed* it seems most likely that it was the effect of the preceding l that prevented the backing of *i. It is also possible, though less likely (see Section 4.5.), that velarisation of *m/___u resulted in a loss of final *u that was earlier than the loss of other final vowels. Thus, at the point at which backing assimilation was to have applied to the form in question, it might no longer have been vowel-final. However, even if this sequence of events were basically correct, a velar consonant could have provided a backing environment in a number of cases. It seems most likely, then, that the preceding consonant was the determining influence.[6]

*iku *tail* became iki- *tail* rather than **uku-. As will be shown below there seems to have been widespread fronting of *u, perhaps to something like [ɨ]. Thus, in this example, the reflex of *u may not have been a back rounded vowel and hence would not have provided a backing environment if this latter change had taken place before backing assimilation of *i. In the case of *pitu?u *star*, the development of a w-glide from *? (see Section 5) would have produced a backing environment; thus uju *star*. It is not clear whether we are to interpret the influence of [w] as having been to prevent the fronting of *u or, at some later period, of 'rebacking' a previously fronted [ɨ]. That this process in fact might be synchronic is suggested by the fact that a more narrow phonetic transcription of uju *star* might resemble [yjy(w)] (the final [w] is questionable), where the quality of the vowels might be attributable to the presence of j, but might also reflect their status as phonemic i. This analysis gains further support from alternations like [eij] *seven, serial* and [yjyw] *seven, general* < /iji+wV/. The two solutions are roughly as follows:

| *pitu?u | | *pitu?u | |
|---|---|---|---|
| yijuwu | Glide-Formation | yijuwu | Glide-Formation |
| — | *u-fronting | yijɨwɨ | *u-fronting |
| ijuu | Glide-deletion | ijiwi | Glide-deletion (and reinterpretation of i) |
| ujuu | u-assimilation | ijiw | Final V Deletion |
| uju | Final V Deletion | yjyw | Rounding Assimilation |

A much finer phonetic study of the primary MOK data will be necessary to evaluate the second solution.

*i > o

This reflex is attested in two forms:

| | | | | |
|---|---|---|---|---|
| *pakiwak | *shark* | > | pako | *shark* |
| *nsiwa | *nine* | > | to- | *nine* |

Synchronically to- alternates with tu and tiwɔ- in the following forms:

toopas, toomɛn, tookij *nine — long object, animate, partitive*
tiwɔw [tuwɔw] *nine — general*
atu [atiw] *nine — serial*

The to- alternate occurs preconsonantally in this paradigm, in the same environment that yields an o reflex of *-iwa- in the case of *pakiwak. Although admittedly based on only two forms, this fact allows us to present an interesting hypothesis as to their derivation:[7]

| **pakiwak | | **nsiwa+pas[8] | |
|---|---|---|---|
| pakewak | | tewa+pas | Lowering |
| pakook | | too +pas | Vowel Hiatus |
| pakoo | | -------- | Final Consonant Deletion |
| pako | | -------- | Final Vowel Deletion |
| pako | *shark* | toopas | *nine, long object* |

Thus a sequence *iwa becomes [oo] */___C.

*i > e

| | | | | |
|---|---|---|---|---|
| *mpu(dr)it | *excrement* | > | pwirej | *dirt* |

I cannot explain the lowering of *i in this item, although the fact that the final consonant is preserved suggests a deleted suffix which might have contributed to the lowering (perhaps underlying **pwirij-a).

*i > Ø

This reflex occurs inexplicably in *inum *drink* > nim *drink*.[9]

## 4.5. *u

*u > u

MOK u is seldom fully rounded thoughout its articulation but tends, rather, to be influenced by its consonantal environment. It retains its rounding only when long and in a few monosyllables closed by rounded or 'back' consonants.

| | | | | |
|---|---|---|---|---|
| *tu?u(d) | *stand up* | > | uuta | *stand up* |
| *puko | *net* | > | uk | *net* |
| *puku | *knot* | > | pwuk | *knot* |
| *upu | *young leaves* | > | upw | *young coconut* |
| *mpua | *areca nut* | > | pwu | *betel nut* |
| *Ruŋma | *house* | > | umw | *house* (imwaa *his house*) |
| *?umu | *oven* | > | umw | *oven* |
| *mudi | *behind* | > | mwuur | *afterwards* |

| | | | | |
|---|---|---|---|---|
| *muta? | *vomit* | > | mwwuj | *vomit* |
| *mputo | *navel* | > | pwujjɔ | *his navel* |
| *punti | *banana* | > | wus | *banana* |
| *?usa | *rain* | > | wut | *rain* |
| *tuki | *pound..* | > | juk | *pound* |

In most of the examples above *u yields [u] in a closed syllable containing a velar or velarised consonant.[10] The closed syllable restriction accounts for alternations like umw *house* and imwaa *his house*.

In several of the cases *u went to u where a *C was lost, possibly leaving in its place a glide homorganic with the following *V.

| | | |
|---|---|---|
| *tu?u+taV | *puko | |
| wuwu+taV | wuko | Glide Formation |
| uu+taV | uko | Glide Deletion |
| uu+ta | uk | Final Vowel Deletion |
| uuta | *stand up* uk | *net* |

uuta *stand up* is derived with the suffix -ta *up* to account for the length of the vowel. If that length can be shown to be a synchronic juncture phenomenon, then the base for *stand* could be derived as u-. It seems likely that the vowel following *u in these cases must be [+rnd], in order to account for cases like:

| | |
|---|---|
| *puki | |
| wuki | Glide Formation |
| wɨki | *u-fronting |
| wiki | ɨ-reinterpretation |
| ---- | Glide Deletion (inapplicable) |
| wik | Final Vowel Deletion |
| wik | *change* |

In all these cases where Glide Formation is a factor (*tu?u(d), *puko, and *puki etc.), the issue of whether *u-fronting never applied or was later reversed by a 'back rounded' environment remains problematic (see Section 4.4.). Similarly, *o in *puko *net* might have blocked *u-fronting or, alternatively, might have 'velarised' *k before Final Vowel Deletion, thus providing a 'rounding' environment.[11] No matter how its role is to be interpreted in the final solution, it remains clear that the sequence *ko is relevant to the 'preservation' of rounding in a preceding *u. (The rule of Glide Deletion will be discussed at greater length below.)

More problematic cases of an u reflex of *u include wus *banana* < *punti and wut *rain* < *?usa as compared with win *scales, feathers* < *?una and wij *pull out* < *putik. I have no explanation for the rounding in these cases. Even stranger is juk *pound* < *tuki, where a rounding environment is even less in evidence. Finally, I must rely for the moment on grade alternation to explain why *p was not lost in *puku > pwuk *knot*. I do not find this explanation satisfying, however.

*u > ɨ

This seems to be the more common reflex of *u, found in all environments except those outlined above.

| | | | | |
|---|---|---|---|---|
| *dua | *two* | > | riɔw | *two, general* |
| *inum | *drink* | > | nim | *drink* |
| *ʔuluŋa | *pillow* | > | wiliŋ | *rest one's head* |
| *sulu | *torch* | > | til | *torch* |
| *kuRita | *octopus* | > | kiij | *octopus* |
| *tapu- | *conch* | > | jowi | *conch* |

Problematic cases include:

*tumpu(ʔ) *grandparent.. > ipwiipw *parentage*, where we would have expected **upwuupw, and

*namu(k) *mosquito* > amwijɛ *mosquito*

The most likely explanation for this latter form is that the i is epenthetic.

## 5. GLIDE FORMATION

As pointed out above (Section 3.7.), *n, p, ʔ, R, s, and t all have zero as one of their reflexes. Let us consider the hypothesis that they were replaced by a glide agreeing in backness and rounding with the following *V (the rule of Glide Formation used in several derivations above) rather than being lost without a trace. This is proposed in order to explain cases like the following:

*p > w

| | | | | |
|---|---|---|---|---|
| *tapu- | *conch* | > | jowi | *conch* |
| *puki | *return* | > | wik | *change* |
| *punti | *banana* | > | wus | *banana* |
| *tapu(n)i | *afterbirth* | > | jɔwjɔw | *afterbirth*[12] |

*p > y

| | | | |
|---|---|---|---|
| *api | *fire* | > | ɔy |
| *kopu(t) | *mist.. | > | sɔɔkoi |

*ʔ > w

| | | | | |
|---|---|---|---|---|
| *ʔuna | *scales* | > | win | *feathers, scales* |
| *ʔuluʔa | *pillow.. | > | wiliŋ | *rest one's head* |
| *ʔusa | *rain* | > | wut | *rain* |
| *ʔuta | *inland* | > | ɛwij | *inland* |
| *ʔuda(ŋ) | *lobster* | > | wuronna | *lobster* |

There are, of course, other examples of the same environments where no glide appears:

*p > Ø

| | | | | |
|---|---|---|---|---|
| *napo | *surf* | > | no | *wave* |
| *puko | *net* | > | uk | *net* |
| *pitu | *seven* | > | ɭjiw | *seven*, general cl. |

*ʔ > Ø

| | | | | |
|---|---|---|---|---|
| *ʔumu | *oven* | > | umw | *oven* |

These last examples point to the existence of a rule deleting a glide that agrees with a following vowel in backness and rounding (the rule of Glide Deletion mentioned earlier).

It would be reasonable to suppose that a theory permitting the development of y and w might also allow an h-glide, as in Bender's solution for MRS. Some examples do in fact suggest a phonetic [h]: ɔj *liver* ∿ hɔj, but al *road* % **hal. This might turn out to be some sort of onset phenomenon, however.

*n > Ø

| | | | |
|---|---|---|---|
| *namu(k) | *mosquito* | amwijɛ | *mosquito* |
| *pan(i) | *wing* | pɔ | *arm, wing* |
| *ponu | *turtle* | woy | *turtle* |
| *tan(ou)m | *bury* | jɔujɔu | *grave* |

*ponu > woy *turtle* seems to suggest that *n > Ø after the application of *u-fronting (to explain the front glide). This conclusion is partially substantiated by the fact that MRS wẹn, Kiribati (Gilbertese) (KIR) on, and TRK wiin *turtle* still show n. The form jɔujɔu, however, seems to be a counter-example. Note that, in this case, even if *n-loss had occurred before *u-fronting, the resulting reflex would still be exceptional in the same way as that of *tapu(n)i, discussed in Note 12. The similarity between these two reflexes invites further study.

*p > Ø

For *napo > no *wave* and *mapo > mo *heal*, I hypothesise the following development:

| *napo | *mapo | |
|---|---|---|
| nawo | mawo | Glide-Formation |
| nɔwo | mɔwo | Raising 1 |
| nowo | mowo | Raising 2[13] |
| noo | moo | Glide Deletion |
| no | mo | Final Vowel Deletion |

no *wave* mo *heal*

KIR seems to reflect these forms without Raising and Final Vowel Deletion as nao *wave* and mao *heal*.

Forms like jowi *conch* < *tapu and wik *change* < *puki suggest that glide formation preceded the fronting of *u, while sɔɔkoi *fog* *kopu(t) suggests the opposite. It is possible that the latter is not the best of cognates, however. In wus *banana* < *punti and woy *turtle* < *ponu, w is preserved in an environment where we would have expected it to have been deleted. wus, however, also contains an anomalous u reflex of *u. The fact that a word poses two problems is a solution to neither, but in this case it does suggest that we not consider wus a strong counterexample to Glide Deletion until we have solved the problem of the vowel reflex. As suggested in Section 4.3., woy *turtle* may not be a counterexample if its underlying form is taken to be /wey/. This also serves to substantiate the rule ordering in the developments of *pitu?u posited in Section 4.4., where Glide Deletion was ordered before Rounding Assimilation.

A final problem concerns *api > ɔy *fire*. In the rule sequence proposed so far, this form should have the following derivation:

| | |
|---|---|
| *api | |
| ayi | Glide Formation |
| ɔyi | Raising |
| ɔi | Glide Deletion |
| **ɔ | Final Vowel Deletion |

The fact that the result is a non-occurring form suggests that either Glide-Deletion or Final Vowel Deletion has not applied. Since this is my only example of *...pi#, it remains a problem.[14]

*? > Ø

This environment shows the expected developments in all but two cases; wut *rain* < *?usa and wuronna *lobster* < *?uda(ŋ). As with wus *banana*, wut shows an unexpected u reflex of *u. In the case of wuronna (if it is a good cognate) the o reflex of *a, which is itself strange, may account for a somewhat more recent backing of i to u.

*R > Ø

No glide is reflected in forms exhibiting this change, even where we might expect a glide, given examples like *api > ɔy *fire*. For example: *paRi *ray* > pɔ *ray* and nsakaRu *reef* > tɔkɔ *island*.

*s > Ø

| | | | | |
|---|---|---|---|---|
| *sala(n) | *road* | > | al | *road* |
| *tasimi | *sharpen* | > | jaim | *sharpen* |
| *sida | *they* | > | iir | *they* |
| *suli | *shoot..* | > | il | *small plant* |

If il is a cognate, it suggests that loss of *s (like loss of *n) is a more recent development than loss of *p and *?.

The treatment of glide formation outlined above was to account for the following correspondences:

*u > wi, in cases like *puki > wik

{*p, *?} > w, in cases like *ponu > woy and ?usa > wut

*a > o, in cases like *napo > no

The last case is the most abstract in that no surface glide appears. An alternative solution does not come readily to mind.

For the first case it would be possible to construct a theory in which wi < [ɨ] (a rounded central reflex of *u in Pre-MOK) by a rule of prevocalic glide insertion and concommitant unrounding and fronting of [ɨ] (perhaps in terms of a contradictory feature solution like that proposed for Eng. vowels in Krohn 1971). Such a theory might provide an explanation for alternations like wina]$_{tr}$ and winaun]$_{int}$ *to scale, pluck*. Whether this solution would, in fact, be very different from the one proposed above remains to be seen. An alternate solution for the second correspondence would involve a rule like *p > w / $\underline{\quad}$ V $\underline{\quad}$ , with subsequent deletion of w in some as yet undetermined environment.

All these alternatives are in the direction of decreased generality in the treatment of glides. The rules by which glides are formed might be

similar in both treatments, except that in the former they are treated as a single phenomenon, while in the latter any similarity between individual rules of glide formation is coincidental. In any event, w seems to be the only glide to appear regularly. At the moment [y] and [h] are theoretical constructs yet to be justified on a synchronic basis.

## 6. OUTLOOK

Very little more progress can be made in the history of MOK without a better understanding of the synchronic phonology. In this regard two areas seem salient: 1) paradigm alternations and 2) vowel/glide phonetics. The first is important in that it might provide insight into the general direction of change in MOK and also a set of rules that might mirror historical change. The phonetic area seems to hold the most promise. Interest in MC vowels, both from a synchronic and diachronic point of view has proceeded through Dyen's work on TRK and Bender's 'provocative' solution for MRS. I feel that an important contribution can be made in trying to reconcile Bender's four-vowel three-consonant-series solution with the facts of other MC languages.

At present I feel that this kind of solution is not suitable for a synchronic grammar of MOK, principally because it lacks the relatively large number of consonant contrasts found in MRS. Thus, MOK jik *tern* and juk *pound*, and ik *tail* and uk *net* are minimal pairs distinguished by their syllabic nuclei, while Marshallese węq *net* and yęk *fish* differ with respect to the consonants surrounding phonemically identical syllable nuclei. It might still be possible to analyse MOK ik and uk with differing initial glides, but as yet there seems to be little justification for setting up more than one velar stop.

Historically, however, I believe it possible to speak of a 'back rounded' environment, as the *ko in *puko *net* > uk *net*, in contrast to a more neutral *kuku > kik *nail*. (I use the term neutral in the sense of unmarked.) The distinction between marked and unmarked environments I hope to formalise to account for the rounding in the example just cited and for the raising and fronting (palatalisation) in mijɔ *his eye* < *mata and jil- *three* < *tolu as compared with the more 'normal', less marked reflexes. It will be interesting to attempt to prove if either the MRS or the MOK pattern is the more conservative or, indeed, if the two systems are part of the same line of development or are merely parallel but unrelated developments.

As has been stressed at numerous points in this paper, it is urgent that we compare the various MC languages with each other. Many interesting problems, and hopefully some solutions, will remain hidden until we are more certain which developments are attributable to PMC and which are language-specific. For example, it has recently come to my attention (Sugita, personal communication) that TRK has a few nouns that must be represented with a long vowel in the base, and that these nouns seem to correspond to a class of MOK nouns with a long stem vowel. (MOK, as I have already mentioned, has no rule of vowel lengthening for monosyllabic nouns.) For example, MOK maan *animal* TRK maan *bird*, MOK rɛɛn *day* TRK råån *day* and MOK pɛɛn *bait* TRK paa *bait*. It is hoped that further investigation will make it possible to resolve these noun sets in terms of PMC, and that the solution will shed some light on the genesis of the vowel lengthening rule.

## NOTES

1. This paper was written in Spring 1972 for a University of Hawaii graduate seminar on Micronesian linguistics conducted by Professor George W. Grace. The data on which this summary is based were obtained by eliciting from a native speaker of Mokilese (MOK) forms which appeared to be cognate with forms reconstructed for Proto-Oceanic (POC) by Grace and others (see Appendix or Grace 1969, Grace n.d., Blust 1972). Statements about the synchronic phonology of MOK are based on data obtained from the MOK informant. It is not expected that all of the generalisations presented here, both historical and synchronic, will hold up in the light of further work on the comparison of Micronesian (MC) languages and on the system of MOK itself.

2. Two asterisks (**) preceding a form indicate that it is nonoccurring. A single asterisk is used to mark reconstructed proto-forms.

3. Polysyllabic forms showing retention of a POC final vowel are:

| | | | | |
|---|---|---|---|---|
| kimi | *we, excl* | < | *kami | *we, excl* |
| kimwi | *you, pl* | < | *kamu | *you, pl* |
| japi | *bowl* | < | *tampi | *bowl* |
| jowi | *conch* | < | *tapu- | *conch* |
| kɔw | *fish-hook* | < | *kau | *fish-hook* |

*tapu- *conch* is not exceptional if the hyphen may be taken to represent an unidentified *V. In all but the last case (where the reconstruction *kau might be wrong for Micronesia — cf. MRS kayaj, PNP kɛɛs) all these forms end in i preceded by a reflex of ...*a$\left[\begin{smallmatrix} C \\ +\text{labial} \end{smallmatrix}\right]$. The only other example of this environment, *api *fire* > ɔy *fire* might also be interpreted as reflecting *i# (see the discussion of this example in Section 5). I hesitate to draw any conclusion from these observations, however.

4. The prenasalisation in nsa may reflect an earlier stage of initial gemination, as evidenced in TRK ccaa *blood* (cf. Rehg on nasal substitution, in this volume).

5. Following Dyen, I include my discussion of *R among the palatals and apicals. Dempwolff suggests a more velar articulation for this proto-phoneme.

6. Note that this argument is not trivial even if the vowel in question can be shown to change its backness and rounding during its articulation, since synchronic u can be shown to be acoustically very different from this vowel. The argument becomes trivial only if u is phonotactically impossible in this environment. I am not certain whether this is in fact so.

7. The actual mechanics of the rules in question are unclear. They must precede final consonant deletion, a fact which may reflect their age, depending on the theoretical position one adopts concerning rule ordering

and reordering. What is probably at issue is whether final consonant deletion is a synchronic rule of MOK grammar, was one at the time these rules were introduced, or had not entered the grammar of MOK at that time.

8. The underlying form of pas might be given as **pasV, depending on the synchronic/diachronic position of Final Consonant Deletion.

9. Note also the apparent retention of the final consonant. This form is aberrant in almost all of the MC languages, reflecting something like PMC *nima.

10. The distinction I am making is between those consonants reconstructed as *mp and *ŋm and those occurrences of *p and *m that yield pw and mw.

11. The fact that this k is not rounded synchronically is, quite obviously, a problem to this 'Marshallese' type of analysis.

12. Note that the last example, *tapu(n)ɨ, appears to be a counterexample to the claim made about o reflexes of *a (in Section 4.1.10) following glide formation. I have no explanation for this particular case.

13. In view of the counterexamples cited above, the exact nature of Raising 2 is still very much in doubt. An alternative would be welcome but it is not yet clear to me what that might be.

14. In this regard it is significant that if uju *star* < *pitu?u is analysed synchronically as /ijiw/ (see the developments posited in Section 4.4.), then the whole analysis of glide formation presented in this paper will have to be re-examined.

## APPENDIX

### Preliminary list of Mokilese cognates

The POC reconstructions used here are from three sources. Those found in Grace (1969) appear here with no special designation. Those based on Dyen 1949 (cf. Grace n.d.b) are preceded by the letter D and those from Blust (1972) by the letter B.

| | POC | | MOK | |
|---|---|---|---|---|
| 1. | *-ʔaki | *cause or instrument* | -ki | instrument suffix |
| 2. | *alap | *take* | ɔlɔ | *take* |
| | | | aljapw | *divide land* |
| 3. | *aŋoaŋo | *yellow* | ɔŋ | *tumeric bush* |
| 4. | *ʔapaRa | *shoulder* | aprɔ | *his shoulder* |
| 5. | *api | *fire* | ɔy | *fire* |
| 6. | *ansa(n) | *name* | at | *name* ɔtɔ *his name* |
| 7. | *ʔate | *liver* | ɔj | *liver* ɔjɔ *his .....* |
| 8. | *ʔatop | *thatch* | ɔj | *thatch* |
| 9. | *awaŋ | *mouth* | aw | *mouth* ɔwɔ *his mouth* |
| 10. | *-da | *their* | -ra | *their* |
| 11. | *daʔa(n) | *branch* | ra | *branch* raa *its* |
| 12. | *ndaŋma | *forehead,...* | sɔmwɛ | *his forehead* |
| 13. | *ndamu | *chew betel* | samwsamw]$_{int}$ samwɛ]$_{tr}$ | *lick* |
| 14. | *(dr)an(i) | *day* | rɛɛn | *day* |
| 15. | *daRa(ʔ) | *blood* | nsa | *blood* nsaa *his...* |
| 16. | *nrau(n) | *leaf* | sɔ | *leaf* sɔɔ *its...* |
| 17. | *deŋa | *tumeric, yellow* | rɔŋrɔŋ | *yellow* |
| 18. | *doŋo | *hear* | roŋ | *hear* |
| 19. | *(dr)on(st)o | *night* | ros | *darkness* |
| 20. | *dua | *two* | ari | *two – serial* |
| | | | riɔw | *two – general* |
| 21. | *ika(n) | *fish* | ik- | *prefix on some fish names* |
| 22. | *iku | *tail* | ik | *tail* |
| 23. | *inu(m) | *drink* | nim | *drink* |
| 24. | *ka(dr)u | *scrape* (*karis) | karpwit | *scrape with hands* |
| 25. | *ka(Rdr)aŋpa | *new* | kapw | *new* |
| 26. | *ŋkala | *male genitals* | kɔlɔ | *his penis* |
| 27. | *kami | *we, excl* | kama | *we, dual excl* |
| | | | kimi | *we, collective excl* |
| 28. | *kamu | *you, pl* | kamwa | *you, dual* |
| | | | kamwi | *you, collective* |
| 29. | *kani | *eat, food* | kanaa | *his food* |
| 30. | *kau | *fish-hook* (*kawa) | kaw | *fish-hook* |
| 31. | *kau | *tree* (*kai) | suukɔ | *tree* |
| 32. | *kaputu | *stomach* | (ʔ)kapeet | *stomach* |

| | POC | | MOK | |
|---|---|---|---|---|
| 33. | *kianto | *outrigger boom* | kia | *outrigger boom* |
| 34. | *kiekie | *pandanus* | kiɔ | *cl., his mat* |
| | | | kipar | *pandanus* |
| 35. | *kima | *giant clam* | kim | *kind of clam* |
| 36. | *ŋkinit | *pinch,...* | kini | *pinch* |
| 37. | *kinta | *we, incl* | kisa | *we, incl dual* |
| | | | kiis | *we, incl collective* |
| 38. | *ko(e) | *you, sg* | kɔɔ | *you, sg* |
| | | | (∿ kɔwɔ) | |
| 39. | *kopu(t) | *mist, dust* | sɔɔkoi | *fog* |
| 40. | B*ŋkompu | *viscera* | (?)pwuk | *bladder* |
| 41. | *kuku | *claw, nail,...* | kik | *nail* kikin *his...* |
| 42. | *kuli(t) | *skin, bark* | kil | *skin, bark* kilin *his...* |
| 43. | *kuRita | *octopus* | kiij | *octopus* |
| 44. | *kutu | *louse* | pakit | *delouse* |
| 45. | *laki | *great,...* | lɛklɛkin | *big* |
| 46. | *lako | *go* | la | *away* |
| | | | (?)lakɔn | absolute superlative marker |
| 47. | *laman | *sea, lake* | lam | *lagoon* |
| 48. | *laŋi(t) | *sky* | lɔŋ | *sky* |
| 49. | *laŋo | *fly (insect)* | lɔŋ | *fly* |
| 50. | *lanje | *branching coral,...* | (?)lar | *branching coral* |
| 51. | B*lato | *stinging nettle* | (?)lɔj | *leader of a fishing party* |
| 52. | B*(ln)ima | *bail out* | lim | *bailer* |
| | | | limɛɛla | *bail out* |
| 53. | *lima | *five* | alim | *five, serial* |
| | | | limɔw | *five, general* |
| 54. | *limu | *seaweed* | limw | *seaweed* |
| 55. | *lo | *in* | lɔlɔ | *inside* |
| 56. | *loku | *bend* | lok | *change direction, crooked* |
| 57. | B*lompi | *fold* | limlim | *fold* |
| 58. | *mada | *fermented* | mar | *preserved breadfruit* |
| 59. | *mai | *come* | mɛ | *here* |
| 60. | *malino | *calm,...* | (?)mulun | *calm* |
| 61. | *maliŋ | *sour* | (?)mɛɛn | *bitter* |
| 62. | *ŋmalo | *submerge,...* | (?)mwoiti | *sunk* |
| 63. | *malu | *soft* | (?)mo | *grass* |
| 64. | *-mami | *our, excl* | -maa | *our, dual excl* |
| 65. | *manaŋ | *power, wind* | manman | *spiritually powerful* |
| 66. | *manipi(s) | *thin* | minipnip | *thin* |
| 67. | *mapo | *heal* | mo | *heal* |
| 68. | *masaki(t) | *pain* | mɔtɔk | *pain* |
| 69. | *masawa | *sea,...* | mataw | *open sea* |
| 70. | *mata | *eye* | maj | *eye* mijɔ *his...* |
| 71. | * mata | *worm* | mwaj | *worm* |
| 72. | *mataku(t) | *afraid* | mijik | *afraid* |
| 73. | *manu(k) | *animal* | maan | *animal* |
| 74. | *mansu(rR) | *plenty of food* | mɔt | *fat* |
| 75. | *mate | *die* | meeti | *die (of animates)* |
| | | | mɛɛta | *die (of plants)* |
| 76. | *matolu | *thick* | mojul | *thick* |

| | POC | | MOK | |
|---|---|---|---|---|
| 77. | *matudu(R) | *sleep* | mɔyr | *sleep* |
| 78. | *ma?udi(p) | *life,...* | mour | *life* |
| 79. | *-mu | *your, sg* | -mw | *your, sg* |
| 80. | *mu?a | *front* | mwoɔ | *in front of it* |
| | | | mwoon | *in front of* |
| 81. | *mudi | *behind* | mwirin | *behind, after* |
| | | | mwuur | *afterwards* |
| 82. | *muta(?) | *vomit* | mwwuj | *vomit* |
| 83. | *-na | *his, her, its* | -n | *his, her, its* |
| 84. | *namu(k) | *mosquito* | amwijɛ | *mosquito* |
| 85. | *napo | *surf* | no | *wave* |
| 86. | *natu | *child* | nɔy | *my child* |
| | | | naa | *his child* |
| | | | niin | *child of* |
| 87. | *niu(R) | *coconut* | ni | *coconut tree* |
| 88. | *nua | *inside* | nɛɛn | *inside* |
| 89. | *onom | *six* | oon | *six, serial* |
| | | | wonow | *six, general* |
| 90. | *pada | *pandanus* | kipar | *pandanus* |
| 91. | *pakiwak | *shark* | pako | *shark* |
| 92. | D*pali | *holiness, prayer* | (?)pɔl | *related so as to exclude marriage* |
| 93. | *pani (*mpaya) | *bait* | pɛɛn | *bait* |
| 94. | *pan(i) | *wing* | pɔ | *arm, wing* pɔɔ *his...* |
| 95. | *paŋ(ou)(n) | *awaken, arouse* | paŋin]$_{tr}$ | *call* |
| | | | pɔŋpɔŋ]$_{int}$ | |
| 96. | *mpampa(n) | *board* | pɔp | *board* |
| 97. | *paRi | *stingray* | pɔ | *ray* |
| 98. | *pat | *four* | ɔpɔŋ | *four, serial* |
| | | | paaw | *four, general* |
| 99. | *mpeŋka | *flying fox, bat* | pwɛɛk | *bat* |
| 100. | *pi(dr)i | *fold, twist* | kapir | *twist* |
| 101. | *pili(?) | *select, choose* | pil | *choose* |
| 102. | *pinsiko | *flesh* | utuk | *flesh* utukin *his...* |
| 103. | *pitu | *seven* | eij | *seven, serial* |
| | | | ijiw | *seven, general* |
| 104. | *pitu?u | *star* | uju | *star* |
| 105. | *mpo- | *smell* | pwo | *smell* |
| 106. | B*mponot | *blocked* | pwon | *blocked* |
| 107. | *ponu | *turtle* | woy | *turtle* |
| 108. | *ŋpoŋi | *night* | pwoŋ | *night* |
| 109. | *mpou | *post* | pwo | *pole* pwoɔ *his...* |
| 110. | *pua(?) | *fruit* | waa | *its fruit* |
| | | | wɛɛn | *fruit of* |
| 111. | *mpua | *areca nut* | pwu | *betel nut* |
| 112. | *mpu(dr)i(t) | *excrement* | pwirej | *dirt* |
| 113. | *puki | *return* | wik | *turn, change* |
| 114. | *puko | *net* | uk | *net* |
| 115. | *puku | *knot* | pwuk | *knot* |

| | POC | | MOK | |
|---|---|---|---|---|
| 116. | *mpulu(t) | *gum,...* | pwilij | *sticky* |
| | | | pwil | *gum, sap* |
| 117. | *puti | *banana* | wus | *banana* |
| 118. | B*putik | *pluck, pull out* | wij | *pull out* |
| | | | (?)wijik | *carry* |
| 119. | *mputo | *navel* | pwijjɔ | *his navel* |
| 120. | *Ripa | *go close* | ip- | *near, close to* |
| 121. | *Ruŋma(?) | *house* | umw | *house* imwaa *his...* |
| 122. | *sala(n) | *road* | al | *road* |
| 123. | *nsaŋi | *wind* | ɛŋ | *wind* |
| 124. | *saŋka(?) | *hop,...* | takɔ | *drive a vehicle* |
| 125. | *nsakaRu | *reef,...* | tɔkɔ | *island* |
| 126. | *nsake | *upwards* | ta | *up* |
| 127. | *nsama | *outrigger* | tam | *outrigger* |
| 128. | *sampo | *catch* | tapɔr]$_{tr}$ | *catch* |
| | | | taptap]$_{int}$ | |
| 129. | *sida | *they* | iir | *they* |
| 130. | *sili | *enter* | til | *enter* |
| 131. | *nsipo | *downwards* | ti | *down* |
| 132. | *nsiwa | *nine* | atu | *nine, serial* |
| | | | tuwɔw | *nine, general* |
| 133. | *nsoka | *stab* | tɔkɔ]$_{tr}$ | *stab* |
| | | | tok]$_{int}$ | |
| 134. | *solo | *mountain* | tol | *mountain* |
| 135. | B*suli | *shoot, sucker* | il | *small plant* |
| 136. | *sulu | *torch* | til | *torch* |
| 137. | *suRi | *bone* | si | *bone* siin *his...* |
| 138. | *-nta | *our, incl* | -saa | *our, dual incl* |
| 139. | *taku | *back* | likinjarkin | *his back* |
| 140. | *tali | *rope* | jɔl | *rope* |
| 141. | *tama | *father* | jamaa | *his father* |
| 142. | *tan(ou)m | *plant, bury* | jɔwjɔw | *grave* |
| 143. | *taŋi(s) | *weep* | jɔŋ | *weep* |
| 144. | *tampi | *bowl* | japi | *bowl* |
| 145. | *tapu- | *conch* | jowi | *conch* (? *older* jɔwi) |
| 146. | B*tapu(n)i | *afterbirth* | jɔw | *afterbirth* |
| 147. | *tansi(k) | *salt, salt water* | jɛt | *salt, salt water* |
| 148. | D*tasimi | *sharpen* | jaim]$_{tr}$ | *sharpen* |
| | | | jɔyjɔy]$_{int}$ | |
| 149. | *ta?u | *year, season* | joonpar | *year* |
| 150. | *tido | *look at* | iroŋ | *look at* |
| 151. | *tina | *mother* | inaa | *his mother* |
| 152. | *tolu | *three* | ejil | *three, serial* |
| | | | jiliw | *three, general* |
| 153. | *tuki | *hammer, pound* | juk | *pound* |
| 154. | *tunu | *cook,...* | iniin | *cook* |
| 155. | *tumpu(?) | *grow, grandparent* | ipwiipw | *parentage* |
| 156. | *tuRi(a) | *to string* | ir | *to string* |

| | POC | | MOK | |
|---|---|---|---|---|
| 157. | *tuʔu(d) | *stand up* | uuta | *stand* |
| 158. | *ʔuda(ŋ) | *lobster* | wuronna | *lobster* |
| 159. | *ʔuluŋa | *pillow,...* | wiliŋ | *rest one's head* |
| 160. | *ʔumu | *earth oven* | umw | *earth oven* |
| 161. | *ʔuna(p) | *fish scale* | wina | *to scale* |
| | | | win | *feather, hair,* winɛ *its feather* |
| 162. | *ʔunsa(n) | *rain* | wut | *rain* |
| 163. | *ʔuta(n) | *inland, forest* | ɛẃij | *inland* |
| 164. | B*upu | *young leaves* | upw | *unripe coconut* |
| 165. | *walu | *eight* | awaꜜ | *eight, serial* |
| | | | waliw | *eight, general* |
| 166. | *wao | *forest woods* | wɔwɔ | *forest* |

## BIBLIOGRAPHY


BENDER, Byron W.

1968 Marshallese phonology. *Oceanic Linguistics* 7/1:16-35.

1969 Vowel dissimilation in Marshallese. *Working Papers in Linguistics, University of Hawaii* 1/1:88-96.

BLUST, Robert A.

1972 Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list. *Working Papers in Linguistics, University of Hawaii* 4/1:1-43.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

1965 *A sketch of Trukese grammar*. Essay 4. New Haven: American Oriental Society.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

n.d. Some Trukese cognates (from Dyen 1949). Mimeo. University of Hawaii.

KROHN, Robert

1971 On the sequencing of tautosegmental features. *Working Papers in Linguistics, University of Hawaii* 3/5:17-25.

# SEGMENTAL QUANTITY IN MOKILESE: A SYNCHRONIC AND DIACHRONIC STUDY

Sheldon P. Harrison

## 1. INTRODUCTION[1]

Segmental quantity distinctions are a feature of most Oceanic (OC) languages[2] but have not been reconstructed for either Proto-Oceanic (POC) or Proto-Austronesian (PAN). The present study does not attempt to demonstrate that extant POC and PAN reconstructions are inadequate in this respect, though this may eventually prove true. Its purpose is to give a history to the segmental quantity distinctions found in one OC language, Mokilese (MOK), a Micronesian (MC) language of the eastern Caroline Islands. While some claims regarding segmental quantity distinctions in Proto-Micronesian (PMC) and in putative lower order subgroups of MC arise from this study, the main focus remains the history of MOK.

Some attempt will be made to provide justification for the major quantity-generating processes considered. The conclusions that follow are admittedly weak, on the one hand because of the lack of a solid theoretical framework and, on the other, because of the paucity of secure comparative data. The study is presented in this rather unpolished form, nonetheless, in the hope that it will act as a stimulus for further research in this area.

Section 2 presents the phonemic inventory of MOK. The origin of vowel length distinctions is considered in Section 3: through the morphological processes of reduplication (3.1.) and affixation (3.2.), through assimilation of vowel sequences (3.2.), and through consonant loss (3.3.). Section 3.4. investigates the possible role of a bimoric canonical target in the genesis of vowel length. Section 4 investigates the history of consonant gemination in MOK. The pre-enclitic lengthening process termed juncture lengthening is considered in Section 5. Section 6 provides a short summary of conclusions reached.

## 2. SEGMENTAL INVENTORY OF MOK

### 2.1. Consonants

The following chart gives the contrastive consonants of MOK:

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 375-401. *Pacific Linguistics*, C-80, 1984.

| | Bilabial | Dental | Palatal | Velar | Velarised labial |
|---|---|---|---|---|---|
| Stop | p | d | j | k | pw |
| Nasal | m | n | | ng | mw |
| Fricative | | s | | | |
| Lateral | | l | | | |
| Trill | | r | | | |

The symbols used are orthographic. Stops are typically voiceless and lenis. All consonants occur both single and geminate (long) but, in contrast to many other MC languages, geminate consonants are found only medially (inter-vocalically). As will be demonstrated, historical initial and final geminates have been resolved into more 'tolerable' patterns.

Although consonant gemination is contrastive, minimal pairs are rare. The only lexically simplex minimal pair as yet identified is likoau *chapped* and likkoau *cloth, clothing, clothed.*[3] In other extant minimal pairs, one of the members is lexically complex; for example, moange *his head* (third person singular of the possessive paradigm of moang *head*) and moangge *this head,* a combination of moang *head* and the enclitic determiner -e *this.* All MOK enclitics trigger lengthening in the preceding segment of the lexical items to which they are clitic (see Section 5).

## 2.2. Vowels

The following chart gives the contrastive vowels[4] of MOK:

| | Front | Central | Back |
|---|---|---|---|
| High | i | | u |
| Higher mid | e | | o |
| Lower mid | ɛ | | oa |
| Low | | a | |

Minimal pairs for contrastive length can be found for all vowels except /e/:[5]

| | | | |
|---|---|---|---|
| win | *feather* | wihn | *to win* |
| -men /mɛn/ | *one (animate)* | mehn /mɛ̄n/ | *bitter* |
| paj | *nest* | pahj | *hollow of a canoe* |
| ros | *darkness* | rohs | *flower* |
| koap | *to grow* | koahp | *yam* |
| pwur | *to turn* | pwuhr | *to sink* |

# 3. ORIGIN OF DISTINCTIVE VOWEL LENGTH

## 3.1. -VC# and -CV̄# reduplications

The no longer productive process of (C)V(C)V# reduplication discussed in Harrison 1973 was used to derive 'descriptive' verb forms (Harrison:1977) — stative verbs from various word classes and intransitive verbs (Vi) from transitive verbs (Vt), with some semantic restrictions that are irrelevant here. Examples are:

| | | | |
|---|---|---|---|
| pinapin]$_{Vi}$ | *to cover* | pina]$_{Vt}$ | *to cover something* |
| pwirejrej | *dirty* | pwirej | *dirt* |
| poadpoad]$_{Vi}$ | *to plant* | poadok]$_{Vt}$ | *to plant something* |

These derivations involve the operation of three phonological processes; the deletion of final consonants at a relatively early period, the subsequent deletion of final vowels, and a synchronic process of vowel reduction which in some instances appears to have restructured phonological representations.

Vowel reduction applies to forms of the shape #$(C_1)V_1C_2V_2C_3V_3$...; that is, to forms of at least three syllables where the first two syllables are open, deleting $V_2$ or reducing it to [ɨ] if deletion would give rise to an impermissible consonant cluster. It creates alternations like alij *beard*, oaljoa *his beard;* and mijik *afraid*, kamjik *frightening*, kamjiki *to frighten someone*. Reduction does not apply if the first vowel is high and the second non-high, as in the case of pinapin *to cover*, but operates normally if both vowels are high: kurujek]$_{Vi}$ [kurjɛk] *to grate* < kuruj]$_{Vt}$ *to grate something*. This process does not account for the internal vowel loss in forms like pwirejrej *dirty*. In Harrison (1973:67), it is suggested that pairs like pwirej *dirt*, pwirejrej *dirty* were restructured without final vowels during a period when the reduplication in question was still a productive morphological process.

We assume derivations like the following:

| | | |
|---|---|---|
| *pina | poadok[6] | |
| -- | poado | -C# Deletion |
| pinapina | poadopoado | Reduplication |
| pinapin | poadopoad | -V# Deletion |
| -- | poadpoad | Vowel Reduction |

Rightward (C)V(C)V# reduplication yields synchronic long vowels under three sets of circumstances:

i) from roots of the historical shape #VCV#

| | | | | |
|---|---|---|---|---|
| alahl | *striped* | cf. | al | *line* |
| oaroahr | *beach* | < | *oaroa | |
| uruhr | *to laugh* | < | *uru | |
| ilihl | *to steer* | < | *ili | |
| inihn | *to cook* | cf. | inim | *to cook something* |

ii) from roots of the historical shape...(C)V̄#

| | | | | |
|---|---|---|---|---|
| marahra | *light (in weight)* | < | *marā | |
| korohro | *white* | < | *koro | |
| pwiahia | *grey-haired* | cf. | pwia | *grey hair* |
| pwijehje | *ridden with excrement* | cf. | pwije | *excrement* |

dangahnga *lazy* (cf. dangpaur *slacker*, probably a compound built on *dangah, the vowels of which were lost by -V# deletion and vowel reduction.)

iii) from roots of the historical shape -V̄CV

| | | | |
|---|---|---|---|
| pwuhsuhs | *round* | < | *pwuhsu |
| poahjoahj | *smooth* | < | *poahjoa |
| sohrohr | *different* | < | *sohro |

Under the assumption that pre-MOK and probably PMC tolerated only open syllables, the (C)V(C)V# reduplication process copied either a final CVCV, a final VCV in disyllabic roots of this shape, a final long vowel (this counting as two morae) and preceding consonant if any, or the final mora of a penultimate long vowel and a following short syllable.

## 3.2. Affixation and assimilation

As is to be expected, the juxtaposition of a vowel-final prefix and a vowel-initial stem or a vowel-initial suffix and a vowel-final stem gave rise to a long vowel if the vowels were identical or were reconciled by an assimilation process. For example. the causative prefix ka- yields long vowels in forms like kahlua *to lead* (cf. alu *to walk*) and kahre- *to cause* (cf. Kosraean (Kusaiean) (KSR) or∧ *to make*). Other long vowels may be the outcome of historical prefixation where the prefixes ceased to function as such at an early date. This could be true of poahjoahj *smooth* (cf. mahjahj *cleared, not overgrown*) and pahdehd *blurred* (cf. edied *smoky, misty* [probably a borrowing from Ponapean (PNP) — see Note 27]). These examples might involve the causative *pa- reconstructed for POC and the PAN stative marker *ma-, though the paucity of data and the great time depth involved make these observations rather speculative.

It is likely that native speakers no longer identify the ka- of the above examples with the productive prefix ka-, in the latter case because the root *are- is not a lexical item and in the former probably because no form *kahlu or *alua exists. In such cases the lexical relationship between the derived causative and its probable source has become opaque. These contrast with examples like kahdanki *to name* (from adanki *to be named*), which is often pronounced with four distinct syllables [kàadánki]. Thus, it can be inferred that the tendency to merge two adjacent short vowels over a morpheme boundary into a single long vowel increases as the analysability of the item concerned decreases. The single case of merger involving the prefix ja- *not* is morphologically opaque: jehjoa *to not know* (cf. PNP ese *to know*, sehse *to not know*). No examples with ja- parallel to kahdanki are attested.

Suffixation of -ek *intransitive*, -oar *and then*, and -oang *towards* to vowel-final stems yields a long vowel. Thus:

| | |
|---|---|
| piroa]$_{Vt}$ | *to twist something* |
| piroahk]$_{Vi}$ | *to twist* |
| okdekda | *to lie down* |
| okdekdahr | *to lie down and then* |
| kijoula | *to travel* |
| kijoulahng | *to travel onwards* |

In such cases one must assume the operation of a process assimilating the vowel of the suffix to that of the stem. It is not perfectly clear, however, that the observed long vowels did in fact, arise in this manner. This issue will be considered in Section 5.

In the case of jehjoa *to not know* the low level of the prefix ja- *not* has been raised (assimilated) to the following lower mid vowel (probably an original *oajoa *to know*). The result is phonetically [ē], however, raised and fronted

between palatal consonants.[7] Note, however, that assimilation of this sort is not categorical since alongside jehjoa we find cases like kaemen *first (animate)* (from emen *one, another (animate)*). (Cf. also kopwung *to judge* and kapwung *to correct* (from pwung *correct*) and koarjoa *to finish, to empty* and karoji *to empty* (from roj *empty*). Assimilation, then, is also governed by derivational opacity; the more opaque the semantic/formal relationship between a root and an affixal derivative, the more likely assimilation is to take place. (The limiting cases are non-occurrence of the independent root and non-analysability of the affix.)

There is evidence both of diachronic development of long vowels from sequences of non-identical vowels and of synchronic variation between long vowels and sequences of non-identical vocalic elements. Cases of this sort involving assimilation across morpheme boundaries have just been noted. The vowel most frequently derived in this manner is [o̅]. In the data at present available this resulting vowel has two pre-MOK sources: [oɔ] and [ɔi] or [ɔu].[8]

MOK shows synchronic alternations between ooa# and oh- in a small class of inalienably possessed nouns:

| | | | |
|---|---|---|---|
| kooa | *on top of it* | kohn | *on top of* |
| koalooa | *its root* | koalohn | *root of* (koalo *root*) |
| pooa | *on it* | pohn | *on* |
| mwooa | *in front of it* | mwohn | *in front of* |
| pwooa | *his smell* | pwohn | *smell of* (pwo *smell*) |
| pwudooa | *his sweat* | pwudohn | *sweat of* (pwudo *sweat*) |
| sooa | *member of it* | sohn | *member of* |

A final ooa# in the third person singular alternates with oh- in the rest of the paradigm of these nouns. For many speakers, oh- has been extended to the third person singular as well, yielding pwoh *his smell*, pwudoh *his sweat*, etc.

Kiribati (Gilbertese) (KIR) moa *first* and ao *surface*, cognates of MOK mwooa *in front of it* and pooa *on it*, respectively, suggest a PMC *oa# source for MOK ooa# (/oɔ#/) sequences. (KIR ao shows regular loss of PMC *p. I cannot account for the order of the vowels in this form, where the MOK cognate would lead one to expect *oa). POC/PMC short final *a, when protected from final vowel deletion by some suffix later lost, was regularly raised to a lower mid vowel in MOK. The usual outcome was /ɔ/, sometimes /ɛ/ (regularly after a nasal); thus MOK mijoa *his eye* < POC *mata *eye*, MOK koaloa *his penis* < POC *ŋkala *male genitals*, MOK soamwe *his forehead* < POC *ndaŋma *forehead*, kijoa *his gift* (cf. kijakij *gift*). Verified cases of final a from short *a appear to involve relatively recent loss of a final consonant — MOK pina *to cover something*, Marshallese (MRS) pinej *to cover something*, (cf. MOK pine *its cover*), MOK wina *to scale something* < POC *ʔuna(p) *fish scale* (cf. MOK wine *its scale*). Internal raising of *a to /ɔ/ or /ɛ/ under conditions not yet fully understood is also common.

POC *iwa# and *awa# sequences yield *oa# in pre-MOK. Thus: MOK pako *shark*, KIR bakoa *shark* (< POC *pakiwak *shark*, with PMC loss of POC *k#); MOK lawalo *wild* < pre-MOK *lawalawa, jawajo]$_{Vi}$ *to mix* < pre-MOK *jawajawa (cf. jawa]$_{Vt}$ *to mix something*, PNP sewe *to mix something*). We hypothesise that jawa *to mix something* did not develop into *jo because of a later deleted transitive inflection. These cases should yield [o̅] where final vowel deletion

did not apply, as, for example, before the construct suffix PMC *ni: MOK pakohn *shark of*.

Without speculating as to the phonetic motivation of the [oɔ] > [ō] change in MOK, it is likely that the survival of the earlier form in third person singular is due to either of two factors:

i) absolute final position (after loss of the hypothesised third person singular suffix)

ii) generalisation of -oa# (/ɔ#/) as third person singular nominal possessive marking (cf. third person singular examples in the preceding paragraphs)

Cognates of pwooa *his smell* and koalooa *its root* do not show reflexes of the expected PMC *oa#, but rather a final high vowel (glide) — KIR boi and MRS bwii- *smell* (note also Trukese (TRK) pwoen *his smell*); PNP kaleu *root*. As will be seen, it is not uncommon for vowel-glide sequences to yield (or to alternate synchronically with) long vowels in MOK. The nature of the historical development represented by these two examples remains to be clarified.

PMC final *ai and *au yield long vowels in MOK in four attested cases: joh *there is not*, dohluhl *to pass by (of time)*, poah *his hand*, and soah *its leaf* < PMC *t'awu < POC *dau *leaf*. It is likely that joh *there is not* and jaudi *no, it is not*, have a common source, reflected also in KIR tai *do not* via a raising of PMC *ai# ∿ *au# to MOK oau#, later reanalysed to oh#.[9] A similar change, word-internally, is reflected in the MOK doublet dauluhl ∿ dohluhl *to pass by (of time)*. The cognate sets: MOK poah *his arm* (cf. poa *arm*), KIR bai *arm*, MRS pein *his arm* (cf. pā *arm*), TRK pewún *his arm* (cf. paaw *arm*), Ulithian (ULI) pɛɛl *his arm* (cf. pɛɛ *arm*), and MOK soah *its leaf* (cf. soa *leaf*), KIR rau *pandanus thatch*, TRK céén *its leaf* (cf. céé *leaf*), ULI cɛɛl *its leaf* (cf. cɛɛ *leaf*), Sonsorolese (SNS) saawyrA *its leaf* (cf. sááw *leaf*) must have a somewhat different history since the MOK outcome is oah, not oa. The existence of an alternation between a final and a non-final vowel glide sequence in isolate and paradigm forms of these two nouns may have been a contributing factor to the low mid vowel outcome. Note also the pair MOK poaidi *downwind side of the main island of Mokil atoll* (cf. -di *down*) and MOK pahdak *upwind side of....* (cf. -da *up*, MRS -tak *up*). These forms may also reflect PMC *paú *arm*. Their paradigmatic separation from poa *arm* accounts for their independent phonological development. Even if they do have the etymology proposed, an adequate explanation for the long vowel of pahdak *upwind side...* is still lacking.

In four cases, earlier sequences of a high vowel followed by a non-high vowel yield long vowels. Two are based on comparisons with PNP: MOK jahlehl *unable to cope*, PNP sahliel; MOK mahlehl *brain*, PNP mwaliel. (The consonant discrepancy in the latter set may be due to a transcription error). These examples are problematic since a PNP VCiV sequence is normally reflected as $VC_1C_1V$ in MOK (see 4.3.). The final two cases involve the numbers *two* and *nine* (POC *dua *two*, *nsiwa *nine*). These have the following forms in MOK:

| *two* | *nine* | |
|---|---|---|
| ari | adu | serial counting |
| riaw | duoaw | general classifier -w |
| roahmen | dohmen | animate classifier -men |
| rahpas | dohpas | long object classifier -pas |
| riekij | dohkij | partitive classifier -kij |

In view of the earlier observations on the history of POC *iwa# in MC, POC *nsiwa *nine* might be expected to yield MOK *dooa- and KIR *roa *nine* (attested KIR rua). The observed reflexes can be attributed to PMC *sua *nine*, without the *w of the POC reconstruction. Final vowel deletion yields the base -du of the MOK serial counting series. The oa (from PMC *a) of duoaw *nine* cannot be attributed to final raising (a possible explanation if the numeral classifiers are considered to have been independent lexical items fused to the numeral base before the operation of final vowel deletion) if final raising in MOK entered the language after final vowel deletion. The final vowel would then have either been deleted before it could be raised or not have been raised at all. Allowing some other process to account for the raising here, the change *ua or uoa to oh parallels other cases that have been discussed, though again it is difficult to specify precisely the environment in which it took place, except to say that in these cases the oh is not the final syllable of the forms in question.

PMC *r(i,u)a *two* shows a long vowel in two of its MOK reflexes: rahpas and roahmen. (The lower mid vowel of the latter may reflect assimilation to the following lower mid /ɛ/). The form riekij /riekij/ is puzzling, however, in that it does not show a long vowel. This may be due to its higher mid /e/, the result of raising between high segments (see Note 7). Such raising, however, usually involves earlier lower mid vowels: piload *to pick something*, pilepil *to pick*, but pina *to cover something*, pinapin *to cover*. It, therefore, suggests an earlier *rioakij, with a base *rioa- rather than the *ria- suggested by the other paradigm members.

There is some synchronic variation in MOK between vowel-glide sequences and long vowels. For example, dahlimw ∿ daulimw *to wipe*, roahla ∿ roaila (roa *to peel* + i *transitive* + la *away*) *to peel off something*. This can be viewed as a continuation of the same tendencies that gave rise to the historical changes discussed above. This is not to say that it involves a continuation of the same 'rules' in a generative phonological sense. In the view taken here historical processes do not persist as synchronic rules. The pressures to which these processes are a response can persist and surface periodically, but each time they do so they may take a slightly different form in response to the extant phonological system. (This sort of behaviour might be termed a 'historical conspiracy'; the traditional term is 'drift'). The rise of long vowels from rising diphthongs of the sort considered here can be viewed as an abductive innovation (Anderson 1973), sensitivé to the articulatory distance between low or lower mid and high vowels. In such cases, a language acquirer might assume the required bimoric structure to have been attained before the tongue has reached its highest point and thereby reinterpret the product as a long vowel.

## 3.3. Medial consonant loss

POC *? and *R are usually reflected as zero in MC languages (though MOK reflects POC *R as r in some cases — aproa *his shoulder* < POC *?apaRa *shoulder* and ir *to string* < POC *tuRi(a) *to string*). Such loss gave rise to long vowels in forms like rah *its branch* (base form: ra *branch*) < POC *da?u(n) *branch*, ujuhn *star of* (base form: uju *star*) < POC *pitu?u *star*, insah *his blood*[10] (base form: insa *blood*) < POC *daRa *blood*, and kihj *octopus* < POC *kuRita. Medial consonant loss often fed the long vowel generating processes discussed

in the last section; for example, MOK johnpar *year*, KIR tai *time* (from POC *ta?u *year, season*); MOK no *wave*, KIR nao *wave*, MRS ŋo *wave* (from POC *napo *wave*) and MOK mo *healed*, KIR mao *healed* (from POC *mapo *to heal*). The short vowels of the MOK forms are the result of final vowel deletion. One last case involves MOK reflexes of POC *natu *child* (KIR nati *child*, MRS nāji- *child*, TRK nawú- *child*), which is reflected as the MOK stem nih- in all forms of the possessive paradigm except the first and third persons singular (noai *my child* and nah *his child*, respectively). This form seems to have been earlier *nei- (cf. MOK pihrin *envy*, PNP peirin), phonetically *[nei],[11] itself from an even earlier *neji still reflected in the oath mejenneji *on my first born child*). The long vowel of nah *his child* does not seem to come from an earlier *nai since the third person singular form of all MOK possessive classifiers (of which set nah is a member) shows an unexpected -ah# (cf. jamah *his father*, joamoamw *your father;* ah *his thing*, oamw *your thing;* warah *his vehicle*, woaroamw *your vehicle*). It seems to suggest an earlier *na *his child*, whose later history remains a mystery. The failure of the first person singular form to yield a long vowel can be related to its low mid reflex /ɔ/ of PMC *a, in contrast to the hypothesised high mid reflex /e/ in forms that ultimately yielded [ī]. In view of the complex history of MOK vowels it can only be suggested that the different reflexes of PMC *a in this paradigm may be related to the fact that, in the first person singular form, the possessive suffix -i̯ *my* is not independently reflected due to the fact that the stem itself ended in [i̯] (cf. KIR nati *child*, natu (alternate natiu *my child*). Putative pre-MOK forms *nai *my child* and *naimw *your child* — the latter with a suffix, the former effectively without — may have had different phonological histories because of this fact.

## 3.4. Bimoric phenomena[12]

Among the most problematic issues in MC phonology is the existence of a process lengthening $V_1$ in nouns whose historical shape was $*(C_1)V_1(C_2)V$. This process has operated throughout the Trukic (TK) continuum and in PNP. It can be seen in KIR with nouns of the appropriate shape when not accompanied by the article te, but does not appear to have operated in MOK, Pingelapese (PNG), or MRS. Its status in Nauruan (NAU) and KSR is unclear (see Bender 1973:469).

In those languages where final vowels have been deleted, historical $*(C)V_1(C_2)V_2$ nouns are reflected as $(C_1)V_1V_1(C_2)$. $*V_2$ is retained as a voiceless vowel in a number of western TK languages, yielding $(C_1)V_1V_1(C_2)V_2$, and in most cases, as a full vowel in KIR. Lengthened $V_1$ noun forms have short $V_1$ alternates when a possessive suffix follows (and, as already noted, when accompanied by the article te in KIR). Consider the following reflexes of POC *kuli(t) *skin* (STW = Satawalese; PUL = Puluwatese):

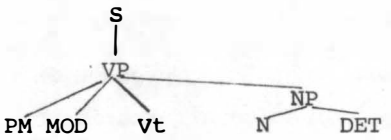
| | Isolate form | 2SG possessed form | |
|---|---|---|---|
| SNS | giirI | girimwU | *your skin* |
| STW | giil | gilimw | " " |
| PUL | kiil | | |
| TRK | siin | sinumw | " " |
| PNP | kihl | kilimw | " " |
| MOK | kil | kilimwen | " " |
| MRS | kil | kiliŋ | " " |
| KIR | kuun, te kun | kunim | " " |

The term compensatory lengthening was applied to this process on the understanding that its motivation was to restore a favoured disyllabic (or at least birmoric) canonical shape to disyllabic forms subject to final vowel deletion. As reported in Bender (1973:467-469), however, evidence suggests that 'compensatory lengthening' preceded loss of final vowels and in fact, applies in languages where final vowels are preserved (as in KIR) or merely devoiced (as in western TK). One might want to view the process as in some sense 'anticipatory', applying initially at a stage when final vowels were devoiced in anticipation of this eventual loss. Its application in KIR remains problematic even under this interpretation, however.

Whatever the motivation of this lengthening process, it is clear that its target, in languages other than KIR, is a bimoric canonical shape (ignoring voiceless final vowels). Thus, it has not applied to otherwise appropriate nouns which show geminate consonants (or reflexes thereof — see Section 4): PNP nta, TRK ccha, Woleaian (WOL) cha *blood*, PNP send *finger*, WOL bun /pʷunn0/ *heart*. It appears that geminate consonants count as one mora.[13]

In PNP, this lengthening does not apply in incorporated object constructions, 'compounds' analogous to English constructions like *to baby-sit* in which the nominal element is directly dominated by the major category V rather than the phrasal category NP. Compare:

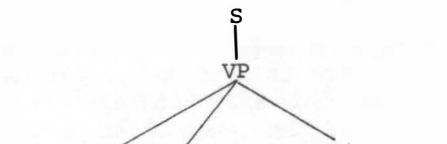
1.




i pahn doakoa mwahmw -o
*I will spear fish that*

I pahn doakoa mwahmwo
*I'll spear that fish.*

2.




i pahn dok mwomw

I pahn dokomwomw
*I'll be spear-fishing*[14]

Thus, a form like dokomwomw *to spear-fish* is treated as a single polysyllabic lexical item in PNP. Similar constructions in TK languages do not seem to have the same analysis (see Sugita 1973:400f).

It has long been a puzzle to Micronesianists that lengthening applies to the major category noun, but not to the major category verb.[15] Two sorts of hypothesis might be used to explain this fact; one syntactic, the other morphological. Under the syntactic hypothesis, lengthening can be viewed as applying to major category items when they occur in initial position in a major category phrase (or, alternatively, to an item exhaustively dominated by a major category). Nouns, then, undergo lengthening either because they are always initial in the NP[16] or because they are exhaustively dominated by N (and no other lexical category). Nouns in PNP incorporated object constructions are not lengthened because they are also dominated by the lexical category V. A similar account may be given for the failure of KIR nouns to lengthen when preceded by the article te.

To demonstrate that verbs cannot undergo compensatory lengthening under this hypothesis one must appeal, for PNP at any rate, to the history of MC syntax. There is evidence that in PMC (and synchronically everywhere in MC except in Ponapeic (PP) and KSR) all predicates were preceded by clitic person markers (often termed predicate markers). This syntactic fact rules out VP-initial position for verbs.

Verbs used attributively (in a NP) typically follow the head noun and hence are not phrase-initial. If they precede the head nouns they usually take a suffix and thus are not monosyllabic. Only the PP languages and KSR do not have obligatory predicate markers. Of these languages, MOK and PNG have no lengthening process, while in KSR all monosyllables are long. In PNP, clitic person markers are required except after noun phrases, but this does not seem to have any effect on the lengthening process in that language.

The major category domination alternative of this hypothesis would involve a claim that clitic person markers are dominated by V, not VP. This would seem to be an unwarranted assumption. Nor can this version of the hypothesis account in any elegant way for the failure of lengthening in attributives.

The morphological hypothesis involves a claim that verbs do in fact undergo lengthening, but of a kind different from nouns. As noted elsewhere (Harrison 1973), monosyllabic verb roots are frequently reduplicated (or undergo initial gemination) when used intransitively (without a transitive suffix) and when exhaustively dominated by V (that is, when, not part of an incorporated object construction).

While this reduplication process does not appear to be motivated by any constraints on canonical shape, since it is found in languages with no lengthening process for nouns and is far from exceptionless, it can nonetheless provide an account of the failure of the lengthening process to apply to verbs. That is, it might be suggested that the very existence of this reduplication process obviates a general rule of lengthening for monosyllabic verbs.

Consider now the following vowel length alternations in MOK mahn *animal*, -men *animate numeral classifier* (from POC *manuk *bird, animal*); mwuhr *afterwards*, mwerin *behind it* (from POC *mudi *behind*); kihs *we (inclusive) as a group*, kisa *we two (inclusive)*, kisai *we (inclusive)* (from POC *kinta *we (inclusive)*); ihr *they as a group*, ira *they two*, irai *they* (from POC *sida *they*).[17] The pronouns ngoahi *I* (KIR ngngai *I*) and ih *he, she* (POC *ia *he, she*) are similarly lengthened, though without short vowel alternates. These items suggest that lengthening did, in fact, apply in MOK and was later lost. The above long forms were 'stranded' when lengthening ceased to operate either because the rule was undone only in monosyllabic nouns (not pronouns or adverbial nouns) or, and

more likely, because only regularly alternating forms were shortened. On the latter hypothesis, we must infer that the lexical relationships holding between the above forms had become opaque enough for their alternatives not to be considered 'regular'.

It might also be suggested that only pronouns underwent compensatory lengthening (perhaps that the process began in the pronouns and later spread to the nouns in the languages west of Mokil). This hypothesis is unappealing, given the present lack of empirical evidence, but cannot be discounted completely, even though it would necessitate a separate explanation for the long vowels of mahn *animal* and mwuhr *afterward* (see below).

Of particular interest are those long monosyllabic nouns whose vowels remain long in polysyllabic paradigm alternates, even in those languages with an active lengthening rule. Although there is not complete agreement among all MC languages as tc which forms follow this pattern, the degree of correspondence is great. MOK has approximately forty long monosyllabic nominals whose length cannot be accounted for. Unfortunately, cognates and/or POC reconstructions can be found for only a few:

| | | |
|---|---|---|
| POC | *manuk | *bird, animal* |
| MOK | mahn | " " |
| PNP | mahn | " " |
| KIR | man | " " |
| MRS | manni | *creature (in compounds)* |
| TRK | maan, ménún | *bird, bird of* |
| ULI | maalI, mann | *bird, animal* |
| | | |
| POC | *(d,r)ani | *day* |
| MOK | rehn | " |
| PNP | rahn | " |
| KIR | rā ni wi | *birthday,* rānnano *daydream*[18] |
| MRS | raan | *day* |
| TRK | ráán, ráánin | *day, day of* |
| WOL | rann | *day* |
| PUL | ráán, ráánin | *day, day of* |
| | | |
| POC | *ŋmane | *male, brother-in-law* |
| MOK | mwahn, mwein | *male* |
| PNP | mwahn, mwan | " |
| KIR | mwmwāne | " |
| MRS | ŋŋaan | " |
| TRK | mwáán, mwáánin | *man, man of* |
| WOL | mwaal, mwanni | " " " |
| | | |
| POC | *pani | *bait* |
| MOK | pehn | " |
| PNP | pahn | " |
| KIR | anea | " |
| MRS | anan | " |
| TRK | paa | " |
| WOL | paa | " |
| SNS | ppa | " |

| | | |
|---|---|---|
| POC | --- | |
| MOK | jahr | *knife* |
| PNP | sahr | " |
| MRS | jāje | " |
| WOL | saar | " |
| | | |
| POC | --- | |
| MOK | rohp | *mat (for sleeping)*[19] |
| ULI | coobw | " " |
| SNS | soobwA | " " |

These data suggest the reconstruction of bimoric roots (trimoric if one counts the final vowel) in PMC for these cognate sets. In most cases a long vowel is suggested, though KIR anea *bait* suggests a final vowel sequence (but see below). The bimoric root is sporadically reflected as a geminate consonant (SNS ppa *bait*, WOL rann *day*), an atypical 'compensatory' response for nouns. These forms are also sporadically treated as short vowel roots (TRK maan, ménún *bird, animal*, PUL ráán, ráánin *day*, KIR man *animal*). The lack of any pattern hints at language-specific restructuring in such cases.

These sets may have undergone lengthening as far back as PMC, so that their reflexes passed into the daughter languages with bimoric root vowels. Three of the six sets can be related to POC forms (with final consonant deletion) ending in a high vowel preceded by a sonorant consonant (POC *manuk *bird, animal*, *(d,r)ani *day*, and *pani *bait*). We might hypothesise that final vowel devoicing may have begun in forms with this canonical shape, yielding long vowels in PMC under the hypothesis that lengthening was 'anticipatory': PMC *rānI *day*, *mānI *bird, animal*, *pānI *bait*. This last reconstruction is problematic because the KIR reflex suggests a final vowel sequence — more evidence is needed here. In another case, PMC *mwāne *male*, a long internal vowel can be reconstructed, but evidence suggests that the final vowel remained intact. (The reconstruction PMC *mwān is possible, except for the KIR reflex). The last two cognate sets are problematic.

If it turns out that some sort of 'anticipatory' lengthening had taken place in PMC, it is interesting to speculate as to whether the lengthening phenomena found synchronically are a generalisation of this process. The geminate consonant reflexes of some of the forms and the fact that all were not restructured with short root vowels in languages with a generalised compensatory lengthening rule for nouns leads me to hypothesise some historical discontinuity, however.

## 4. ORIGIN OF GEMINATE CONSONANTS

The distribution of geminate consonants is not identical in all MC languages, though all, to my knowledge, show such sequences. KIR has initial and medial geminate nasals, but there is no evidence that these have the same history as the geminate consonants of other MC languages. MRS shows initial and medial geminates, the former broken by an epenthetic vowel in the Ratak dialect. So far as I am aware, the only true geminate consonants of PNP are nasal (initial and medial) and lateral (final). Other historical and synchronic geminates have been resolved into homorganic nasal-obstruent sequences in that language. The TK languages show true geminates in all positions, although individual languages may prove to be exceptions to this generalisation.

MOK permits true geminate consonants only medially, although historically it had such consonants in all positions. With the exception of nj#, final consonant clusters are tolerated only in recent Eng. borrowings like pangk *bank*. All historical final geminates (excepting nj# as a possible reflex of older *jj#) are reflected as single consonants in contemporary MOK (see 4.4.). Initial non-nasal geminates have all been resolved, as in PNP, into homorganic nasal-obstruent sequences. Unlike PNP, MOK has an obligatory prothetic vowel in these cases, so that even historical geminate nasals are now no longer initial. Any discussion of consonant gemination in MOK is further complicated by the existence of synchronic processes creating geminates (see 5), which must be distinguished from the reflexes of Proto-Ponapeic (PPP). One must also recognise that relatively recent borrowings from PNP may lead to confusion between direct MOK reflexes of PPP geminates and recent adaptation to canonical shapes not permitted in MOK. Such borrowings are difficult to detect, given the fact that MOK frequently interprets PNP loans in terms of recognised sound correspondences that themselves reflect historical changes.

## 4.1. Initial geminates

As first noted by Goodenough (1963), the initial geminates of MC languages seem to have arisen through the reduction of original #CV- reduplications, as evidenced by a comparison of TRK and KIR cognates:

| POC | TRK | | KIR | | MOK |
|---|---|---|---|---|---|
| *daRa(?) | ccha | | rara | *blood* | insa |
| *muta(?) | mwmwus | | mumuta | *vomit* | umwwuj |
| | kken | | kakang | *sharp* | ingkoang |
| | nnet | *true* | nenera | *search out* | --- |

This change took place only when the original *#$C_1V_1C_1V_1$ shape was the result of morphological reduplication. There are no cases in which fortuitous *#$C_1V_1C_1V_1$ becomes #$C_1C_1V_1$ (for example, MOK poap *board* < POC *papa (n)), nor any cases in which a prefix $C_1V_1$- on a #$C_1V_1$ root yields $C_1C_1V_1$ (for example, MOK kakapwihla *to repair* — ka- *causative* + kapw *new* + -i *transitive* + -la *away*). Gemination from *#CV- reduplication seems to have been further restricted in MOK, not applying to #$C_1V_1$-$C_1V_1V_1C_1$ cases (for example, doadoahk *to work*, mwamwahl *to malign*, koakoahk *tired*), nor when there existed a related unreduplicated form (for example, doadoa]$_{Vi}$ doa]$_{Vt}$ *to sew;* sasa]$_{Vi}$ sahr]$_{Vt}$ *to core*). This last restriction appears to be specific to MOK, as witnessed by the cognate set: TRK ffég]$_{Vi}$ fégúni]$_{Vt}$ *to awaken*, PNP mpeng]$_{Vi}$ pangin]$_{Vt}$ *to call*, MOK poangpoang]$_{Vi}$ pangin]$_{Vt}$ *to call*.

Probable compounds aside, all but four of the twenty-three forms reflecting initial geminates in MOK have the structure [V]$N_1C_1V_1$(C) < PMC *$C_1V_1C_1V_1$(C)V, where [V] is a prothetic vowel and $N_1C_1$ is the homorganic nasal-obstruent reflex of initial gemination. The four exceptions are: imwpwilapwil *pink*, a -CVCV# reduplication of an earlier *imwpwila and, therefore, not a counterexample, ingkarak *large louse* and ingkinik *small louse*, perhaps older compounds, and umwwoais *to play, to be engaged in an activity not related to work*. These three final cases remain counterexamples.[20]

It is possible to construct a hypothesis, albeit quite tentatively, to account for the rise of initial geminates #$C_1C_1V_1$ from earlier #$C_1V_1$-$C_1V$ reduplications based on a constraint that the reduplicated #$C_1V_1$ syllable eschewed primary stress. Though stress assignment in MOK is a complex matter yet to be explored in detail, it appears that primary stress is usually penultimate, though sometimes initial.[21] Given this stress pattern, primary stress in pre-MOK #$C_1V_1$-$C_1V_1(C_2)$, following the loss of final vowels (or their devoicing if voiceless final vowels were not significant for purposes of stress assignment), would have fallen on the reduplicated #$C_1V_1$ syllable in violation of the constraint noted above. The vowel loss giving rise to initial geminates can then be viewed as a mechanism for removing an otherwise stressed syllable so that primary stress would fall on the 'root' vowel. The lack of initial gemination (from #CV reduplication) in KIR might then be considered a result of the retention of final vowels in that language, so that primary stress remains on the penultimate 'root' vowel. Further, the fact that #CV reduplication is retained in MOK before a long vowel (for example, doadoahk *to work*) is explained by the fact that primary stress falls on the long 'root' vowel, where stress is seen as sensitive to vocalic moras rather than to syllables as such. The very tentative nature of these proposals must be borne in mind.

## 4.2. Final geminates

The existence of final geminate consonants at some earlier period in the history of MOK can only be inferred indirectly through PNP. The only surface final geminate consonant found in contemporary PP languages is PNP ll#, as in mall *field* (MOK mal *cleared area*) and dill *penetrate* (MOK dil). Given the fact that all geminate obstruents are realised in PNP as homorganic nasal-obstruent clusters (PNP nda *to say*, nta *blood*, doadoahnki (MOK doadoahkki) *to work with* [PNP/MOK doadoahk *to work* + -ki *with*] one might infer that PNP $N_1C_1$# in such native lexical items as mand *tame*, send *finger*, emp *coconut crab*, ndand *famous*, and engk *mountain* reflect earlier final geminates in similar fashion. The evidence in favour of this proposal is not unequivocal, however, since cognate items are both few and difficult to evaluate. One such cognate set is:

| | | |
|---|---|---|
| PNP | mand | *tame* |
| MOK | maid | " |
| KIR | manana | " |

If the KIR reflex is assumed to mirror the PMC shape of this form, one can postulate the following history for PNP mand:

| | | |
|---|---|---|
| PMC | *mánana | |
| | manna | Reduction of post-stress vowel |
| | mann | Final Vowel Deletion |
| | mand | Nasal Geminate Dissimilation |

This case represents the only evidence to date for the dissimilation of final nasal geminates, if the KIR form is in fact cognate, and if its shape cannot be otherwise explained. The postulation of a process deleting post-stress vowels (under circumstances as yet unclear) in the history of PNP final nasal-obstruent clusters is appealing, but requires more substantive evidence. The

case in question, moreover, points twoards at least some instances of initial stress in PMC.

All synchronic $N_1C_1$# clusters in PNP cannot be accounted for in terms of earlier $N_1N_1$# geminates. Since PNP cannot be assumed to have ever had more than the four contrastive nasals (m, n, ng, and mw), which could give rise to mp#, nd#, ngk#, and mwpw# (the last unattested), the sequences ns# (PNP kens *ulcer*) and ns# (PNP kent *urine*) are left unexplained. Other cognate sets are more suggestive of final geminate non-nasals. e.g. PNP emp, omp *coconut crab* MOK opup *coconut crab*. In the absence of conflicting evidence, the mp# of the PNP forms can be assumed to represent a development from two labial stops as reflected in the MOK cognate. What is not completely clear, though, is whether the u of the MOK form was deleted at some stage in the history of the Ponapean forms (yielding *opp, later omp ∿ emp), or is to be considered an epenthetic vowel inserted to break up an earlier final geminate in MOK. Non-PP cognates might resolve this issue.

Consider now the following cognates for PNP send *finger:*

| | | |
|---|---|---|
| PNP | send, sendi | *finger, his finger* |
| MOK | jaid, jandin | " " |
| KIR | tati | *claw* |
| MRS | janit, jantin, | *finger, his finger; claw, its claw* |
| SNS | gaatY ∿ gattY | *finger* |
| WOL | gatt | " |
| TRK | éwút | " |

The TK cognates are problematic since they reflect PMC *#k, rather than the PMC *#t suggested by the other forms. The cognacy of the KIR form is doubtful in view of the fact that MOK d is usually reflected as KIR r. The TK forms, if cognate, suggest that the n appearing in the PP and MRS reflexes should not be reconstructed for PMC and make it likely that the i of the MRS and ú of the TRK isolate forms are epenthetic. A final cognate set, PNP pwand *slow, late* MRS bat *slow* bbat *slow, late* is problematic. Though the data are admittedly scant, it can be tentatively concluded that at least some cases of PNP $N_1C_1$# clusters derive from earlier final geminate obstruents. The same would be true for MOK where, though only one NC# cluster nj# is attested in native vocabulary, alternations like jaid *finger* jandin *his finger* suggest that others can be reconstructed for some earlier stage of the language.

## 4.3. Medial geminates

All MOK consonants can be found geminated medially:

| | |
|---|---|
| oppop | *to pull* |
| pwiddihdi | *wart* |
| wajejjej | *to skip (over water)* |
| likkoau | *cloth* |
| jipwwu | *to tip (in baseball)* |
| pijimmere | *coconut meat* |
| wuronna | *lobster* |
| kanggir | *belt* |
| jumwwa | *chicken* |
| wahssa | *red* |
| pallan | *to compete against* |
| liporro | *seed* |

Some medial geminates are to be found in Jap. borrowings — jipwwu *to tip (in baseball)*, nappa *Chinese cabbage*, wannangge *shotput* middo *catcher's mit*. Others are the result of historical or synchronic vowel reduction or assimilation — warra *their (two) vehicle* (from warara), jumwwa *chicken* (cf. archaic jumuwa, where the /m/ of the archaic form has assimilated to the velarity of w yielding synchronic mww, wahssa *red* (cf. PNP wehtata *red*),[22] poappa *the board* (poap *board* + -wa *the*), pilelle *clarified coconut oil* (pilen *water of* + le *coconut oil*), diddil *to stick through* (cf. dilim *to stick through something*). By far the majority of medial geminate consonants are, however, the result of the synchronic process of junctural lengthening, to be considered in Section 5. A residue must be accounted for historically.

In most of these residual cases neither internal nor comparative evidence sheds any light on their genesis:

| | | | | |
|---|---|---|---|---|
| MOK | linnapw | MRS | lennab | *allergy* |
| MOK | pwakke | MRS | bakke | *yaws* |
| MOK | pwiddihdi | MRS | bwitti | *wart* |

Consider, however, the following alternations in MOK:

| | | | |
|---|---|---|---|
| pwiri | *stomach (of a fish)* | pwurroa | *his stomach* |
| pwiki | *knee, cape* | pwukkoa | *his knee* |

These forms suggest earlier *pwurioa *his stomach* and *pwukioa *his knee*, reconstructions confirmed by the following comparative data:

| | | |
|---|---|---|
| MOK | pwukkoa | *his knee* |
| PNP | pwukie | " " |
| MRS | bukwien | " " |
| TRK | púkúwen | " " |
| MOK | pallali | *counterpart* |
| PNP | peliali | " |
| MRS | pāleen | *his spouse*, pālele *married* |
| TRK | pwúnúwen[23] | " " |
| MOK | pallan | *to compete against* |
| PNP | pelian | " " " |
| MOK | kirrej | *hawkfish* |
| MRS | kidiej | " |
| MOK | kirred | *muddy* |
| PNP | kiried | " |
| MOK | jallali | *meat course, sauce* |
| MRS | jālele | " " " |
| MOK | joalloa | *his meat course, sauce* |
| MRS | jāleen | " " " " |

On the basis of these data it can be hypothesised that at least some instances of medial gemination in MOK derive from earlier VCiV. This hypothesis receives additional confirmation from the reduplicated intransitive form of verbs whose historical root seems to have been of the form #VCV#, for example:

| | | |
|---|---|---|
| oppop]$_{Vi}$ | oapi]$^{24}_{Vt}$ | *to pull* |
| onggong]$_{Vi}$ | ungud]$_{Vt}$ | *to wring out* |

| | | |
|---|---|---|
| edded]$_{Vi}$ | idip]$_{Vt}$ | *to draw water* |
| oarroar]$_{Vi}$ | oari]$_{Vt}$ | *to dismantle (an earth oven)* |
| oalloal]$_{Vi}$ | oaloa]$_{Vt}$ | *to take* |

These pairs can be derived (ignoring vowel alternations) via the historical processes of final consonant deletion, -(C)v(C)V# reduplication, final vowel deletion, and a medial gemination rule $V_1C_1V_2V_1 \Rightarrow V_1C_1C_1V_1$ (where $V_2$ is higher than $V_1$) (see below):

| | | |
|---|---|---|
| *ongud | ongud+V | |
| ongu | --- | Final Consonant Deletion |
| onguongu | --- | Reduplication |
| onguong | ongud (> ungud) | Final Vowel Deletion |
| onggong | | Gemination |

Unless the synchronic vowels of oaloa]$_{Vt}$ *to take something* were of different heights at some historical period one would expect the reduplicated intransitive form to be *oaloahl (cf. al *road, line* alahl *lined*). The PNP cognate ale *to take something* suggests this might have been the case. Though both these forms can be related to POC *alap *to take* (cf. KIR anaa *to take something*), it is sufficient that the vowels of the MOK reflex were of different heights at the point when reduplication took place to guarantee a geminate outcome.

The suffix -ek in MOK regularly affixes to transitive verbs to derive intransitives: kidim]$_{Vt}$ kidimek]$_{Vi}$ *to wrap*, poadok]$_{Vt}$ poadokek]$_{Vi}$ *to plant*, pihni]$_{Vt}$ pihniek]$_{Vi}$ *to paint*. Four verbs show gemination before -ek:

| | |
|---|---|
| sipis]$_{Vt}$ | *to tie something* |
| sippek]$_{Vi}$ | *to tie* |
| lukum ∿ lim]$_{Vt}$ | *to fold something* |
| limmek]$_{Vi}$ | *to fold* |
| weij ∿ wij]$_{Vt}$ | *to pull up something* |
| wijjek]$_{Vi}$ | *to pull up* |
| pid]$_{Vt}$ | *to wind something* |
| piddek]$_{Vi}$ | *to wind* |

In the first example, the -ek intransitive is formed not on the transitive verb but on what appears to be the transitive verb root, (termed non-transitive form in Harrison 1977) used in incorporated object constructions (sip joal *to tie rope*), with gemination of the final consonant. In the next two cases, the -ek intransitive is formed on the 'short' transitive (of CiC shape) rather than on the more common longer transitive. (These 'short' transitives may be back formations on the -ek intransitive. The KIR nukuma ∿ rukuma *to fold something* suggests a PMC *lukuma *to fold something*). In these three examples, then, the root of the -ek intransitive appears to have been haplologised. KIR evidence suggests a similar history for MOK pid]$_{Vt}$ piddek]$_{Vi}$ *to wind*, KIR biroa *to twist something*, whose MOK cognate we would expect to be *pidoa]$_{Vt}$. These four cases of gemination can be viewed as a response to haplological phenomena which, at present, defy systematic explanation.[25]

It was demonstrated above that the cases of MOK medial geminates under consideration here (with the possible exception of the -ek intransitive examples) derive from earlier $V_1C_1V_2V_1$ sequences in which $V_2$ was higher than $V_1$. This structure is not reflected, however, in a number of non-PP cognates.

| | | | |
|---|---|---|---|
| MOK | pwiri, | pwurroa | *(fish) stomach, his stomach (emotionally)* |
| KIR | buro, | burona | *heart, his heart* |
| MRS | bōro, | būruon | " " " |
| ULI | bwull, | bwullal | " " " |
| POC | *mputo | | *navel* |
| MOK | pwijjoa | | *his navel* |
| KIR | buto, | butona | *navel, his navel* |
| MRS | bwijen | | *his navel* |
| SNS | bwuut0, | bwuutorA | *navel, his navel* |
| TRK | bwuuw, | bwuuwen[26] | " " " |

MRS bōro is a regular cognate of KIR burō *heart*. If PMC *pwurō is reconstructed the MRS paradigm forms (būruon *his heart*, etc.) suggest a 'breaking' of the PMC final *ō, ultimately giving rise to the TK/PP forms. No other examples of this phenomenon are extant, however. Both the KIR and MRS reflexes of POC *mputo *navel* are regular, while the TK/PP forms suggest innovations in the direction of a bimoric root (a long vowel in TK and an earlier *#CVCiV# structure in MOK). The motivation for this change is opaque, however. A most tentative hypothesis might relate it to the inherent weakness of high vowels that we see in MOK where, for example, high vowels are often 'skipped over' by processes like vowel reduction (see 3.1.) and stress placement; for example, áriar *gaff*,[27] not *arí́ar, mijoá *his eye*, not *mijoá́. One notes that, in a large number of the forms with medial geminates discussed here (the -ek intransitive cases and many of the inalienable nouns), the root vowel is high. Moreover, a surprising number of these forms also show adjacent high consonants (palatals, velars, and velarised labials). While these facts are suggestive, their explanatory value remains weak.

The medial gemination process $V_1C_1V_2V_1 \Rightarrow V_1C_1C_1V_1$ (where $V_2$ is higher than $V_1$ hypothesised to govern the cases of medial gemination discussed above can be viewed as another instance of the drift toward a closed syllable pattern in MOK (a previously noted instance being the process of vowel reduction — see Section 3.1.). The maximum number of closed syllables in a structure like pre-MOK *opiop (later oppop *to pull*) can be obtained by adjusting the syllable structure from o.pi.op to op.i.op. The resulting internal $i$ would then tend to become $i̯$, yielding the structure op.i̯op with two closed syllables. It is to be expected, however, that glides in structures of this shape will be unstable, being interpreted as vowel colouring on the preceding consonant as in the Eng. rapid speech rendering of [gat.yə] *got you* as [gǽ.čə]. Such a reinterpretation in the MOK structures at issue (op.i̯op > o.pʸop) would have the result of creating an open syllable from a closed one. A closed syllable pattern could, however, be maintained if the $C^G$ of these $VC^GVC$ structures was reinterpreted as a geminate so that it functioed at the same time as coda for a preceding syllable and onset for a following. It is hypothesised that maximisation of the number of closed syllables in this way provided the motivation for the change from pre-MOK $V_1C_1V_2V_1C_1$ to $V_1C_1C_1V_1C_1$.

## 4.4. Dissolution of geminate consonants

The history of MOK reflects moves towards the dissolution of initial and final geminate consonants. Medial geminates appear to be stable and, in fact, to some degree favoured, there being at least two synchronic processes giving rise to such geminates. One of these processes, junctural lengthening, will be treated in Section 5. The second involves assimilations of three sorts:

i) of liquids to obstruents

sassalip *to be applying pomade* (from salip *to apply pomade*)
diddil *to stick through* (from dilim *to stick through something*)
joajjoal *to wind (rope)* (from jal *to wind something*)
joajjoar *to dig* (from jaripw *to dig something*)

ii) of nasals to liquids

illa *to go* (alternate of inla *to go*)
mellau *thanks* (alternate of menlau *thanks*)

iii) of velarisation in labials

poappa *the board* (poap *board* + -wa *the*)
japwwa *the land* (japw *land* + -wa *the*)

This third assimilation type is the only one that crosses lexical boundaries (where noun + wa *the* combinations are noun phrases — see Harrison 1977:208). In this case, the velarity of the initial consonant of the enclitic assimilates to that of the final consonant of the item to which it is clitic. Word-internally, non-velarised assimilates to velarised, as in the case of jumwwa *chicken* (archaic jumuwa).

Assimilations of the first two types appear to follow a consonant strength hierarchy of the sort suggested in Foley (1972) and Zwicky (1972), where obstruents are stronger than nasals, and nasals, in turn, stronger than liquids. The fact that these assimilations are regressive in MOK may, in part, explain the apparent reversal of the principle of assimilation to the stronger of two adjacent consonants in the case of the nasal to liquid assimilations. However, as Foley (1972:97) notes, the assimilation of nasals to liquids in not uncommon (Latin collis *hill*, Lithuanian kalnas) Foley's suggestion that liquids are in fact stronger than nasals does not appear to be completely justified by these facts, since liquid to nasal assimilations are attested. In Yokuts, a Turkic language, the plural suffix -lar becomes -nar after [n]. Note also that in WOL [n] is the tense (geminate) counterpart of [l] (Sohn 1975:15). While the notion 'consonant strength hierarchy' seems to be a useful one, it requires considerably more investigation.

The first stage in the dissolution of initial (and, probably, final) geminate obstruents is their conversion into homorganic nasal-obstruent clusters. This can be viewed as a response to the heightened pressure inherent in geminate obstruents. This pressure can be reduced by lowering the velum to allow some air to escape through the nasal cavity, thereby destroying the obstruent articulation.[28] These reflexes of initial geminates are further weakened in MOK by an obligatory prothetic vowel. This vowel is i unless followed by mw and a round vowel in the next syllable, in which case it is u. Thus: immas *ripe*, insa *blood*, ingkoang *sharp*, umwwoais *to play*, umwwuj *to vomit* (but kaumwwuji *to nauseate*, where a non-syllabic 'prothetic' vowel is always i). The fact that the prothetic vowel appears even after prefixes like ka- 'causative' demonstrates that it is part of the lexical representation of the forms in

question. PNG appears to reflect a further stage in the dissolution of initial geminates, by which they are reflected as #$\bar{V}$C; for example, PNG ahla *to go* (MOK inla *to go*), PNG ahdoa *to come* (MOK indoa *to come*), PNG ihsa *blood* (MOK insa *blood*). The lack of uniformity of the vowels of the PNG cognates is troublesome. More PNG data are, therefore, necessary in order to correctly evaluate these correspondences.

As noted in Section 4.2., it is less certain that final $N_1C_1$# clusters (or reflexes of such clusters) derive from earlier $C_1C_1$# geminates than it is that initial #$N_1C_1$ clusters derive from earlier initial geminates. Whatever the provenience of such final clusters in PNP, it remains true that they have been altered significantly in MOK. Consider the following correspondences:

| | | |
|---|---|---|
| PNP | ndand | *famous* |
| MOK | indan | " |
| PNP | pwai pwand | *to be in debt* |
| MOK | pwai pwanid | " " " " |
| PNP | sirangk | *trunk* (from Eng. *trunk*) |
| MOK | jirang | " |
| PNP | mand | *tame* |
| MOK | maid | " |
| PNP | send, sendi | *finger, his finger* |
| MOK | jaid, jandin | " " |
| PNP | kens | *ulcer* |
| MOK | keinj | " |

From the above data we see that the weakening of *$N_1C_1$# clusters takes three forms in NOK:

i) deletion of the final obstruent

ii) insertion of an epenthetic vowel

iii) weakening of the nasal to a glide

Since all three manifestations can be seen in the resolution of earlier *nd# clusters, it is safe to assume that they represent independent approaches to the same problem at different historical periods. Since the only examples that show non-PP cognates are of the third type (see Section 4.2.), this can be assumed to be the oldest. The first and second types may represent relatively recent adaptations to PNP borrowings. If jirang *trunk* and palang *porch* are also borrowings from PNP, then they shed no light on the history of any putative *ngk# clusters in earlier periods of MOK. The data available show one example of nj# in non-Eng. derived forms (keinj *ulcer*)[29] and one of a hypothetical *ns# (jais *finger*, an alternate of jaid *finger* in the expression ngapin jais *(a) span*). No forms that might be attributed to earlier nasal-obstruent clusters involving labials occur (but recall MOK opup *coconut crab*). The apparent stability of final nj# is interesting, but without additional data, no further conclusions can be drawn.

Medial geminates, as pointed out above, do not appear to have been weakened. The six examples containing medial $N_1C_1$ clusters known to me (joampoak *unkind*, limpoak *affection*, moadoandoal *smooth*, janjal *clear*, koandoau *ladder*, and

jandin *his finger*) can be otherwise accounted for. The first two cases appear to involve a hypothetical *mpoak and the prefixes ja- *not* and li- *characteristically*. The next two forms show an lC to nC change that is not general in MOK. It is possible that some, or all, of these cases are loans from PNP. The form koandoau *ladder* is possibly a compound involving the verb doau *to climb*, and the last case, as noted in 4.2. reflects an earlier MOK form *jand (extant jaid *finger*). In this example we must assume that the phonological shape of the earlier isolate form was used in the paradigm.

## 5. JUNCTURAL LENGTHENING

Geminate consonants and long vowels arise synchronically in MOK through a general process triggered by a following enclitic. These include the enclitic determiners (both definite and indefinite), verbal directional/aspectual enclitics, the instrumental -ki *with*, -oang *towards* and -jang *from*, and the delimiting enclitics -oar *only, up to/and then* and -oaroh *until*. Any vowel-initial enclitic from this set triggers gemination of a preceding consonant; any consonant-initial enclitic triggers lengthening in a preceding vowel. No adjustment is made to consonant clusters arising through clisis. Vowel-initial enclitics have consonant (or glide) initial sandhi alternates when preceded by a vowel.[30] Some examples of these phenomena are:

I. consonant gemination

1a. woallo *that man* (woal *man* + -o *that*)
b. woal koalikko *that big man*
c. woal koalik roahmenno *those two big men*

2. kadarroang *to send something to* (kadar *to send something*)

3. Ih kin kang rokumwwoar doari. (rokumw *crab*)
*he* MOK *eat crab-only finished*
*He ate nothing but crab.*

II. vowel lengthening

4a. pukohkai
puko wahssahkai

5. Ih pihnihkihla parnij. (pihni *to paint something* + -ki *with* + -la *away*)
*He painted it with varnish.*

6. Penno kupdihjang in kohn niho. (kup *to fall* + -di *down* + -jang *from*)
*That coconut fell from the coconut tree.*

Though usually enclitic to the verb phrase (as defined in Pawley 1972:41), the enclitics ki *with*, oang *towards*, jang *from*, and oaroh *until* can appear outside the verb phrase and separated from it by a preceding noun phrase. In such instances they do not trigger junctural lengthening. Thus:

7. Jerimwein koalikko pokihdi jerimwein siksikko ki suhkoahpas.[31]
*The big boy hit the little boy with a stick.*

8. Ih kihdi ekij mwingehu oang lakapw.
*He set some of that food aside for tomorrow.*

9. Kama sikesikkoang arai oaroh jippo kijoula.

*We chatted with them until the ship left.*

While it is tempting to try to account for these phenomena phonetically either in terms of some feature of the underlying phonological shape of the enclitics or of the forms to which they are enclitic, or historically in terms of the phonetic processes described above as giving rise to segmental quantity distinctions in MOK, such analyses can be shown to be untenable. In the case of consonant gemination, one might want to hypothesise an underlying/historical initial glide for the vowel-initial enclitics. One such enclitic, -oaw *a (general)* does show a glide in its post-vocalic sandhi alternate. Thus:

| | |
|---|---|
| pukkoaw | *a book* |
| pukohioaw | *a basket* |

This is not true of other vowel-initial enclitics, however (pukken *that book (near you)* pukohn *that basket (near you)*, nor could it have been true historically for at least one enclitic, -ok *those*, which seems to have been metathesised from an earlier *-ko (cf. MOK -kai *these*, -kan *those (near you)* PNP -kat *these*, -kan *those (near you)*, -ko *those*). This being the only likely phonetic explanation for consonant gemination before enclitics, its weak motivation suggests that we seek an explanation of some other sort.

Vowel lengthening might be considered to reflect historically (or synchronically underlying) long vowels. This hypothesis can be vitiated by demonstrating that many of the vowels which undergo junctural lengthening were short historically. For example, enclitics like -ki *with* or -da *up* lengthen before other consonant-initial lexical suffixes, as in kijoukihdahla *to travel away eastward (up) by means of*. It is argued in Harrison (1977:186) that -ki *with* reflects PMC *akini, the form of which is better preserved in WOL yagili and KIR akini. It may be that -da *up* can be related to POC *nsake *to climb* (cf. MRS -tak (alternate -ta) *eastward, up* and the MOK nominalised directional -dakoa *it's (V-ing) up*). Any argument that these and similar forms have been restructured with a synchronic underlying long vowel must be abandoned in view of the fact that short vowels often appear in forms built on the same root in cases other than those predicted by the rule of final vowel deletion. For example, the vowel, of si *ear* is lengthened before determiners, as in sihkai *these ears*, but not before the suffixes of the possessive paradigm (sioaioa *my ear*, sioamwen *your ear*, sioa *his ear*, etc.). Similarly, the final vowel of pina]$_{Vt}$ *to cover something* lengthens before directional/aspectual enclitics like -di *down* as in pinahdi *to put a cover on something*, but not in the reduplicated intransitive pinapin]$_{Vi}$ *to cover*. We are again forced to conclude that vowel lengthening before enclitics, like consonant gemination under the same conditions, is not phonetically motivated but conditioned solely by the enclitic status of the immediately following element.

Junctural lengthening before enclitics can be viewed as a direct response to their enclitic status. As MOK moves towards an 'agglutinating' structure where such non-affixal formatives come to form a single phonological word with preceding formatives, junctural lengthening functions to preserve the 'lexical integrity' of individual lexical items by providing a clear surface marking of the boundary between lexical items in an agglutinated string. It prevents resyllabification (as in cases like woalle *this man*, syllabified *woa.le without gemination) and blocks processes like vowel reduction (as in cases like poaloahdi

*to chop something*, surface *poaloadi [pɔldi] without vowel lengthening) which would becloud lexical boundaries. In phonetic terms, junctural lengthening prevents long disfavoured sequences of internal short open syllables which would arise through agglutination. However, its operation seems to be motivated more by lexical than by purely phonological considerations.

## 6. SUMMARY

Two broad mechanisms giving rise to segmental quantity distinctions can be distinguished in the arguments presented here. The first, for the most part restricted to vowel length, involves such uncontroversial phonological and morphological processes as segment loss, assimilation, affixation, and reduplication, all resulting in the juxtaposition of identical vowels and thereby giving rise to long vowels. The second may be characterised in terms of the maintenance of favoured 'target' patterns. Among the targets considered are: bimoric canonical shape, 'root' morpheme primary stress, closed syllables, and phonetically marked lexical boundaries, all of which have been appealed to in this study to account for some cases of vowel length and/or consonant gemination. Teleological accounts of this sort are at present controversial in linguistics, and the present study, because of limited data, cannot claim to have provided strong justification for them. While the results of this study may remain unconvincing, such accounts should, nonetheless, not be dismissed out of hand.

## NOTES

1. This paper was first written in 1974 and only slightly revised in 1978. Since that time a considerable body of data bearing on the issues discussed here has come to light. Considerations of time have prevented a further revision of the paper incorporating the information now available.
2. Vowel length distinctions are widespread in Oceania. Contrastive consonant length is, on available evidence, restricted to MC languages and to the Polynesian outliers within Micronesia.
3. Orthographically, geminate consonants are written double in MOK. The geminate versions of the digraphs pw, mw, and ng are pww, mww, and ngg, respectively.
4. All vowel symbols in this chart are orthographic except /ɛ/, which is not distinguished from /e/ in the orthography: pweipwei [pweipwei] *silly* pweipwei [pwɛipwɛi] *tuna*. The digraph oa represents a vowel whose approximate value is /ɔ/. There is considerable phonetic variation in the realisation of MOK vowels, conditioned for the most part by adjacent consonants. For some discussion, see Harrison, 1976, Section 1.6.
5. Vowel length is indicated by an h following the vowel.

6. Historical final consonants and final vowels remain in transitive verb forms because these are assumed to have taken a vocalic transitive suffix later removed by final vowel deletion.

7. Raising of lower mid vowels to higher mid between high segments can be demonstrated by alternations like piload *to pick something* pilepil /pilepil/ *to pick* and pwuroa *to turn something*, pwuropwur *to turn, to spin.*

8. Possible counterexamples to these developments are not considered here.

9. MC languages frequently show variation in high vowel (or glide) reflexes; for example, MOK ni *coconut tree* KSR nu (from POC *niu(R) *coconut tree.*

10. The history of the consonants in items such as this is treated in detail in Section 4.1.

11. Note that PNG shows the stems neu- (nou-), before [mw], and nei- elsewhere in forms where MOK has nih- in this paradigm.

12. A more recent and thorough treatment of the phenomena considered in this section can be found in Rehg (this volume).

13. WOL ch is a palatal stop, the tense (geminate) counterpart of sh [š]; n is the tense (geminate) counterpart of l. Nasal-obstruent clusters in PNP reflect earlier geminate obstruents — see Section 4.

14. The second vowel of dokomwomw *to spear fish* is excrescent.

15. It has since come to light that imperative forms of verbs of the shape #(C)$V_1$(C)$V_2$# in KIR show a lengthened $V_2$; for example, naako! *Go!* (< nako *to go*).

16. Certain quantifiers like MOK dir *many* can precede the head noun of an NP. It can be argued that these stand outside what might be termed the 'nuclear noun phrase'.

17. Plural pronouns like kisai *we (inclusive)* and irai *they* reflect older trials. The forms ira and irai are archaic, having been replaced, for the most part, by ara and arai.

18. KIR rā is of doubtful cognacy, since POC *d, *r is usually reflected as Ø in KIR and since POC *n is not, to my knowledge, typically lost.

19. There is no obvious explanation for the consonant discrepancy here. It might reflect a transcription error.

20. I discount transitive verb forms like insinge *to write on* (cf. insing *to write*) and umwwujoa *to vomit on* (cf. umwwuj *to vomit*) as true counter-examples, since they reflect the appropriate historical root canonical shape, preserved in these forms before a later deleted vocalic transitive suffix.

21. Rehg (personal communication) has suggested that stress in PNP is realised as a rise in amplitude on the penultimate vowel and of pitch on the antepenultimate. I can neither confirm nor deny a similar pattern in MOK, lacking convincing acoustic data.

22. MOK wahssa *red* might reflect a compound built on the reconstructed pre-MOK *ssa *blood*. I know of no arguments to decide between these two analyses.

23. I have no explanation for the initial consonant discrepancy between the TRK and other cognates.

24. The synchronic vowel alternations shown here are not yet understood. Whatever their provenience, they are irrelevant to the present argument.

25. One might want to attribute these cases, at least in part, to an earlier form of the suffix, *-iek. This alternate appears synchronically, however, in only one form pinaiek *to wind* (from pina *to wind something*). Anyone entertaining this hypothesis would still have to explain the cases in which gemination does not occur, as well as the vowel-final transitives that do not take -iek. I attribute the -i- of the form just cited to recent loss of a final palatal stop in this verb (cf. MRS pinej *to cover*).

26. TK cognates in languages west of Ifaluk seem to reflect an earlier *k, rather than the PMC *t of other cognates. Compare the cognate set for 'finger' in 4.2. above.

27. Forms with extant VCiV sequences in MOK can only be explained as borrowings from PNP.

28. The fact that the dissolution of geminate consonants has not spread to medial geminates in MOK might be explained in part by the fact that the two 'parts' of these geminates are assigned to different syllables.

29. The -i- of MOK keinj *ulcer* might represent a first step in the dissolution of this cluster. It also appears in recent borrowings like dainj *to dance* and Parainj *France*. Note, however, alternations like pwaij *share*, pwoaijoa *his share*, pwajanki *to have as a share*, in forms with single palatal stops. I cannot as yet account for this phenomenon.

30. Note that the consonantal, rather than the vocalic, sandhi alternates of these enclitics appear after high vowels — liho *that woman* (li *woman*), suhkoahu *that tree* (suhkoa *tree*).

31. Sentences with independent -ki do occur, but are not regarded as fully grammatical by all speakers.

## BIBLIOGRAPHY

ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DeBRUM

1976 *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.

ANDERSEN, Henning

1973 Abductive and deductive change. *Language* 49:765-793.

BENDER, Byron W.

1969 *Spoken Marshallese*. Honolulu: University Press of Hawaii.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 426-465. The Hague:Mouton.

1973 Parallelisms in the morphophonemics of several Micronesian languages. *Working Papers in Linguistics, University of Hawaii* 5/8:1-16. Reprinted in *Oceanic Linguistics* 12 (1973):455-477.

BINGHAM, Hiram

1908 *A Gilbertese-English dictionary*. Boston: American Board of Commissioners for Foreign Missions.

1922 *Outlines of a grammar of the Gilbert Islands language*. Rongorongo, Beru, Gilbert Islands: London Missionary Society. (Originally compiled in 1861 in MS form only.)

COWELL, Reid

1851 *The structure of Gilbertese*. Beru, Gilbert Islands: Rongorongo Press.

FOLEY, James

1972 Rule precursors and phonological change by meta-rule. In Robert P. Stockwell and Ronald K.S. Macauley, eds *Linguistic change and generative theory*, 96-100. Bloomington: Indiana University Press.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The Proceedings of the Ninth Pacific Science Congress, 1957*, vol.3:77-86.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

HARRISON, Sheldon P.

1973 Reduplication in Micronesian languages. *Oceanic Linguistics* 12: 407-454.

1977 Some problems in the history of Mokilese morpho-syntax. Doctoral dissertation, Department of Linguistics, University of Hawaii, Honolulu.

HARRISON, Sheldon P. and Salich W. ALBERT

1976 *Mokilese reference grammar*. Honolulu: University Press of Hawaii.

PAWLEY, Andrew W.

1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Readings in Oceanic culture history*, vol.3:1-142. Pacific Anthropological Records 13. Honolulu: Bernice Pauahi Bishop Museum.

QUACKENBUSH, Edward M.

n.d. Comparative Trukic word list.

QUACKENBUSH, Hiroko C.

1970 Studies in the phonology of some Trukic dialects. Doctoral dissertation, University of Michigan, Ann Arbor.

SABATIER, Ernest, M.S.C.

1971 *Gilbertese-English dictionary*. (Translated by Sr Mary Oliva, F.N.D.S.C., of Tarawa, Gilbert Islands, Sacred Heart Mission.) Sydney: South Pacific Commission Publications Bureau.

SOHN, Ho-min

1975 *Woleaian reference grammar*. Honolulu: University Press of Hawaii.

SOHN, Ho-min and Byron W. BENDER

1973 *A Ulithian grammar. Pacific Linguistics*, C-27.

SUGITA, Hiroshi

1973 Semitransitive verbs and object incorporation in Micronesian languages. *Oceanic Linguistics* 12:393-406.

WELLEY, Weldis and Elaine GOOD

1976 A short grammar of Pingilapese. MS.

ZWICKY, Arnold

1972 Notes on a phonological hierarchy in English. In Robert P. Stockwell and Ronald K.S. Macauley, eds *Linguistic change and generative theory*, 275-301. Bloomington: Indiana University Press.

Harrison, S.P. "Segmental quantity in Mokilese: a synchronic and diachronic study". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:375-401**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.375*

# KOSRAEAN REFLEXES OF PROTO-OCEANIC PHONEMES

Kee-dong Lee and Judith W. Wang

## 1. INTRODUCTION

This discussion of Kosraean (Kusaiean) (KSR) reflexes of the phonemes reconstructed for Proto-Oceanic (POC) is based on a set of approximately 100 putative cognates culled by Kee-dong Lee[1] from some 180 KSR forms possibly cognate to the POC reconstructions given in Grace 1969, Dyen 1949, and Blust 1972. The purpose of this paper is to present the major correspondences between present-day KSR segments and the phonemes reconstructed for POC, as represented in these data, and to consider how the one set might have developed from the other.

Let us begin with a brief discussion of the sound system of KSR. Lee 1975 gives the following inventory of distinctive surface vowels:[2]

| | front | central | back |
|---|---|---|---|
| high | i [i] | ih [ɨ] | u [u] |
| higher mid | e [e] | uc [ə] | o [o] |
| lower mid | ac [ɛ] | uh [ʌ] | oh [ɔ] |
| low | ah [æ] | a [ɑ] | oa [o͡ɑ] |

The KSR standard orthography adopted in 1973 clearly is not phonemic; however, it will be used in this paper in the absence of an adequate analysis of the KSR phonological system. The phonetic values given for the vowels are those in Lee 1975.

Like several other Micronesian[3] (MC) languages, KSR has three different sets of consonants: plain, velarised, and labialised (phonetically rounded). The following table of KSR consonants is adapted from Lee 1975.

| | | bilabial | labiodental | dental | palatal (retroflex) | velar |
|---|---|---|---|---|---|---|
| stops | plain | p | | t | | k |
| | velarised | pw | | tw | | kw |
| | labialised | | | to | | ko |
| frica- | plain | | f | s | sr | |
| tives | velarised | | fw | sw | srw | |
| | labialised | | | so | sro | |

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 403-442. *Pacific Linguistics*, C-80, 1984.

| | | | | | |
|---|---|---|---|---|---|
| nasals | plain | m | n | | ng |
| | velarised | mw | nw | | ngw |
| | labialised | | no | | ngo |
| liquids[4] | plain | | l | r | |
| | velarised | | lw | rw | |
| | labialised | | lo | ro | |

As with the vowels, the symbols in this table are those of the 1973 standard orthography.

Since the sequence of consonant followed by w is relatively rare in KSR, a w following a consonant generally can be interpreted as a diacritic. A following o is used to represent distinctive rounding — non-distinctive rounding is not spelled out. With one or two exceptions, both distinctive velarisation and distinctive rounding appear to be restricted to consonants preceding a mid front vowel (e or ac). Non-distinctive velarisation of consonants before central and back vowels and non-distinctive rounding of consonants before round vowels are not indicated in the standard orthography. All three types of consonants appear to occur non-distinctively in syllable-final position.

KSR also has two glides w and y. W occurs both velarised and rounded (not distinguished orthographically), but never plain. Y has all three variants, of which only the rounded y is distinguished orthographically as yo.

KSR monosyllables generally conform to the canonical shape (C)V:(C). In citation forms, at least, all KSR monosyllables have long vowels — this appears to hold for lexical (content) words in context, but is not necessarily true of grammatical morphemes. Disyllables may have one of a number of shapes, all of which can be collapsed into the schema (C1)V1(C2)(C3)V2(C4). V1 may be either long or short, but V2 appears always to be long. One- and two-syllable words are the most common in KSR; however, words of up to five (underlying) syllables are possible — e.g., sefolosuwohs *(adj) sleeping lightly*. In normal speech many of these longer words will lose a syllable, even two, through processes of vowel reduction and loss. It should be noted, however, that consonant clusters are permitted only in medial position. Such clusters frequently, though not always, represent the juxtaposition of two morphemes. Thus, the basic morpheme (as well as (mono-)syllable) shape in KSR appears to be (C)V(C).

## 2. CONSONANT CORRESPONDENCES

The table below summarises the consonant correspondences represented in the data on which this paper is based. Conditioning environments are not given for the different reflexes in the table because it is not always clear what the conditioning factors might have been. In many cases different reflexes of the same proto-segment occur in the same environment, suggesting either sporadic change or possibly dual inheritance.

| | | | | |
|---|---|---|---|---|
| POC | *p | *t | *k | *ʔ |
| KSR | Ø,p,y | t,s,n[5] | k | Ø |
| POC | *mp | *nt | *ŋk | |
| KSR | f,p | s,sr,t | k | |
| POC | | *s | | |
| KSR | | s,sr,t,Ø | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| POC | | | *ns | | | | | |
| KSR | | | s,t,Ø | | | | | |
| POC | *m | | *n | | *ŋ | | | *ŋm |
| KSR | m,w | | n | | ŋ | | | w,m |
| POC | | *l | | *d[6] | *(dr) | | *R[7] | |
| KSR | | l | | l | r | | Ø | |
| POC | | | | *nd | | | | |
| KSR | | | | sr | | | | |
| POC | | *y | | | | | *w | |
| KSR | | Ø | | | | | w,Ø | |

## 2.1. Stops

### 2.1.1. *p

POC *p is normally lost in KSR:

| POC | | KSR | |
|---|---|---|---|
| *papine | | acn | *woman* |
| *pitu | | it- | *seven* |
| *punti | | usr | *banana* |
| *topu | | tuh | *sugarcane* |
| *napo | *surf* | noa | *wave* |
| *pai- | | a- | reciprocal pref |
| *paka- | | ahk- | causative pref |
| *patu | | yot | *stone* |

In the last example above it appears as though *p > y. Why this should have occurred is not clear. Ordinarily one would expect to find the glide agreeing with the vowel in backness and roundness, regardless of whether one believes that the consonant was lost first and an epenthetic glide inserted, or that the glide was inserted first and then the consonant lost, or even that the glide resulted from the weakening of the consonant and is thereby a direct successor of the consonant. It is possible that *p was a palatalised segment, in this item at least, and that the y somehow preserves the palatality of the lost consonant. In that case, the glide must have arisen before the consonant was lost, but it would appear not to be a direct successor by weakening of the consonant, since [y] is not on the most natural path from [p] to Ø via [b] > [β] > [w]. The status of y in the synchronic phonology of KSR is unclear. Most of the glides which appear in the standard orthography (both w's and y's) appear to be phonemic. It is possible that there was a diachronic process of glide insertion in the development of KSR from POC. Whatever the case, a single instance of *p > y is not much basis for any kind of generalisation. It may be worth noting, however, that there is no trace of a glide in the pronunciation of any of the other KSR forms that once contained *p. Only noa *wave* sometimes sounds as though it were [nwa], but the probable explanation for this may lie elsewhere (in the synchronic nature of oa).

The other exceptions to the complete loss of *p are more blatant:

| | | |
|---|---|---|
| *pakiwak | pahko | *shark* |
| *paŋ(ou)(n) | pahngon | *(vt) to call* |

What is not obvious is the factor conditioning the retention of *p in these forms. It is evidently not the presence of a following *a, since *p was lost in forms like *panua *land* (> KSR acn), as well as in the reciprocal and causative prefixes (see above) and the troublesome *patu (> yot). What it is, is not evident.

A further question occurs to us: if POC *p normally is lost in KSR and *mp normally becomes KSR f (see Section 2.2.1.), then where does KSR p come from? p is not at all rare in contemporary KSR, but its historical source at the POC level is not clear. Further investigation will be required to unravel this mystery.

### 2.1.2. *t

POC *t is most frequently reflected as KSR t.

| | | |
|---|---|---|
| *tama | tuhma- | *father* |
| *taŋi(s) | tuhng | *to cry* |
| *tolu | tol- | *three* |
| *tu?u | tu | *to stand* |
| *pitu | it- | *seven* |
| *natu | nahtuh-/nuhtV-[8] | *child* |
| *?utup | uti | *(vt) to draw water* |
| *muta(?) | wihte | *(vt) to vomit* |
| *matolu | mahtol | *thick* |

However, when followed by *i or *e, *t appears to become KSR s.

| | | |
|---|---|---|
| *?ate | acsyac-/esyac- | *liver* |
| *tia(n) | (in-)siyac- | *belly* |
| *ŋkinit | kihnis | *(vt) to pinch* |
| *mpulut | fulus | *(vt) to paste* |

The change of *t to KSR s in the last two examples above can be accounted for by supposing the presence of a transitivising suffix *-i on the verb (see Harrison 1978).

In one case *t is reflected by KSR s in an environment in which one would normally expect to find KSR t: *tali > sucl *rope*.

The KSR word ninac for *mother* is probably cognate to the POC form *tina, with an idiosyncratic change in the initial segment.

### 2.1.3. *k

The KSR reflex of *k in non-final position is k, as shown in the following examples:

| | | | |
|---|---|---|---|
| *kianto | | kiyacs | *outrigger boom* |
| *kinta | | kuht | *we (1 pl)* |
| *tuki | | tuk | *(vt) to pound* |
| *kuRita | | koet | *octopus* |
| *kali | | kuhlkuhl | *(vi) to dig* |
| *nsakaRu | *reef* | tuhka | *island* |
| *waka | | okah | *root* |

### 2.1.4. *ʔ

*ʔ appears to have been lost completely in KSR.

| | | | |
|---|---|---|---|
| *maʔudi(p) | | moul | *alive, life* |
| *ʔuluŋa | *pillow, to rest one's head* | ilung | *(vt) to rest, prop, support* |
| *ʔumu | | um | *earth oven* |
| *tuʔu(d) | | tu | *to stand* |
| *ʔate | | acsyac-/esyac- | *liver* |
| *daʔa(n) | | lah | *branch* |
| *tudu(ʔ) | | tul | *drip, drop* |

## 2.2. Prenasalised stops

The following prenasalised stops occur in the POC reconstructions used in this paper: *mp, *nt, *ŋk. The problem of the oral and nasal grades in POC is well beyond the scope of this paper; however, it is worth pointing out that the various OC languages frequently differ not merely on the question of preserving the original oral-nasal grade distinction (supposing there to have been one at all), but also on which grade is reflected in a given lexical item. Some of the confusion in the KSR reflexes of POC prenasalised segments may be due to unresolved confusions at a higher level.

### 2.2.1. *mp

*mp appears to become KSR f with only one exception:

| | | |
|---|---|---|
| *mpeŋka | fak | *bat, flying fox* |
| *mpo- | fo | *smell* |
| *mponot | fonos | *(vt) to block, stop up, plug up* |
| *mpou | fo | *post, pole, stick* |
| *mpoŋi | fong | *night* |
| *mpule | ful | *shell: cowry* |
| *mpulut | fulus | *(vt) to paste* |
| *mputo | fuht/fihtac- | *navel* |

The exception is POC *mpaya > KSR pa *bait*. It is possible that KSR pa comes from some other source such as POC *pani, but then one would have to explain the loss of the entire second syllable of the proto-form. Final vowel deletion could account for the loss of *i (see Section 3.2.), but the loss of *n could only be regarded as idiosyncratic. It is also possible that *mp > p/___*a, although one hesitates to postulate such a rule based on only one instance. Additional data will be needed to decide whether or not this form is an exception to a general change of POC *mp to KSR f.

### 2.2.2. *nt

Only three instances of *nt occur in the data:

| | | |
|---|---|---|
| *kianto | kiyacs | *outrigger boom* |
| *kinta | kuht | *we* |
| *punti | usr | *banana* |

With so little data it is difficult to arrive at any conclusions regarding the regular development of *nt. Both *outrigger boom* and *banana* have alternate reconstructions with *t. KSR s suggests *t in the word for *outrigger boom*, while KSR sr would seem to suggest something other than *t in the word for *banana*.

### 2.2.3. *ŋk

KSR appears to have merged *ŋk with *k, reflecting both as k. This generalisation is based on only two forms, however:

| | | |
|---|---|---|
| *mpeŋka | fak | *bat, flying fox* |
| *ŋkinit | kihnis | *(vt) to pinch* |

## 2.3. Fricatives

### 2.3.1. *s

*s is sometimes preserved as KSR s, as in

| | | |
|---|---|---|
| *sulu | sul | *torch* |
| *sae(t)/*sai/*sei | se | *(vt) to tear* |

In the case of *susu > titi- *breast*, *s appears to have become KSR t under conditions which would appear to favour preservation of s.

*s appears to have been lost in forms in which it preceded *i:

| | | | |
|---|---|---|---|
| *?anusi | | acni | *to spit* |
| *si(dr)i(t) | | iri | *masturbation* |
| *sili | *to enter* | il | *to come, go, move* |
| *tasimi | *sharpen it* | twem | *(vt) to sharpen* |

More data will be needed before anything more definitive can be said concerning developments of POC *s in KSR.

### 2.3.2. *ns

It appears to be the rule that *ns is lost in KSR:

| | | | |
|---|---|---|---|
| *nsake | | -yak/-ack | *up, upwards* |
| *aŋin/*nsaŋi | | eng | *wind* |
| *nsiwa | | yuh | *nine* |
| *tansi(k) | *sea, salt water, salt* | te[9] | *beach, seaside* |

The two KSR words containing a non-zero reflex of *ns show different reflexes in what appears to be the same environment:

| | | | |
|---|---|---|---|
| *nsai | | suc | *who?* |
| *nsakaRu | *reef* | tuhka | *island* |

## 2.4. Nasals

### 2.4.1. *m

POC *m is regularly reflectd as KSR m.

| | | |
|---|---|---|
| *mata | muhta-/mwet[10] | *eye, face* |
| *mapo | mah | *to heal* |
| *matudu(R) | mutul | *to sleep* |
| *tama | tuhma- | *father* |
| *ʔuma | imac | *garden, field* |
| *lima | luhm-/lime- | *five* |
| *tasimi | twem | *(vt) to sharpen* |
| *-mu | -m | 2sg poss suff |

In two instances, however, *m seems to have become KSR w:

| | | |
|---|---|---|
| *manu(k) | won | *bird* |
| *muta(ʔ) | wihte | *(vt) to vomit* |

It is not clear what could have conditioned this change. Not the presence of a following round vowel, because mutul has a (synchronic) following round vowel and *-mu a historical one. *manu(k) has *a following *m, just as in *mata, *mapo, etc., and ih in wihte is not a round vowel. Vowel height cannot be a conditioning factor, for similar kinds of reasons. Nor does it seem likely that a change of *m to w was conditioned by any of the other consonants in either of these forms. The only explanations left to us are sporadic change or borrowing; there is no evidence in the data to show which might have occurred.

### 2.4.2. *n

Non-final *n is normally reflected as KSR n.

| | | |
|---|---|---|
| *manipi(s) | mihnini | *thin* |
| *niu(R) | nu | *coconut* |
| *ŋkinit | kihnis | *(vt) to pinch* |
| *manu(k) | won | *bird* |
| *panua | acn | *land* |
| *natu | nahtuh-/nuhtV- | *child* |
| *inu(m) | nihm | *(vt) to drink* |

### 2.4.3. *ŋ

Non-final *ŋ normally remains ŋ (orthographic ng) in KSR.

| | | |
|---|---|---|
| *laŋi(t) | lucng | *sky* |
| *ŋapulu(ʔ) | si-nguhul | *ten* |
| *ʔuluŋa *pillow* | ilung | *(vt) to rest, prop, support* |
| *paŋ(ou)(n) | pahngon | *(vt) to call* |
| *doŋo | lohng | *to hear* |
| *mpoŋi | fong | *night* |

### 2.4.4. *ŋm

Only two instances of the labiovelar nasal *ŋm occur in the data:

| | | |
|---|---|---|
| *ŋmata | wet/wat | *worm* |
| *Ruŋma(?) | yuwac-/iwac- | *house, shelter* |
| | lohm | *house* |

Disregarding for the moment the form lohm, it appears as though the regular reflex of *ŋm might be w. The form lohm is peculiar in two respects. In the first place, *R is regularly lost in KSR (see Section 2.5.5.). In the second place, the remaining evidence suggests that *ŋm regularly became KSR w rather than m. However, if one were to disregard for a moment one's beliefs about regular reflexes, lohm appears quite plausible as a reflex of *Ruŋma(?). Both the phonetics and the semantics fit well enough. It is possible, though it seems unlikely in view of the meaning and frequency of the word, that lohm is a borrowing and that this is an instance of what Biggs 1965 termed 'indirect inheritance'. The fact that both the initial l and the final m appear to be irregular reflexes of the corresponding proto-segments tends to support this hypothesis. Against it is the fact that no donor language has yet been identified. This, however, may be merely a matter of conducting a systematic search for the likely donors of not only this, but possibly other as yet unidentified loanwords in KSR.

## 2.5. Liquids

As mentioned in Note 6 the term 'liquid' is used loosely in this paper to denote the set of proto-phonemes including *l, *d, *(dr), *nd, and *R. The first three all have KSR liquids l or r as their reflexes. *nd is included because it is the nasal-grade counterpart of *d, although it is not regularly reflected as a liquid in KSR. *R is included here only by virtue of the fact that it appears to correspond to a liquid in some other OC languages.

### 2.5.1. *l

The normal reflex of non-final *l is KSR l.

| | | | |
|---|---|---|---|
| *sulu | | sul | *torch* |
| *walu | | oal- | *eight* |
| *tolu | | tol- | *three* |
| *lima | | luhm-/lime- | *five* |
| *laŋo | | loang | *fly (insect)* |
| *sili | *to enter* | il | *to come, go, move* |

### 2.5.2. *d

KSR l appears to be also the regular reflex of *d.

| | | |
|---|---|---|
| *da?a(n) | lah | *branch* |
| *doŋo | lohng | *to hear* |
| *dua | lo/luo-/lu- | *two* |
| *madama | mahlwem | *moon* |

| | | |
|---|---|---|
| *matudu(R) | mutul | *to sleep* |
| *ma?udi(p) | moul | *alive, life* |
| *tudu(?) | tul | *drip, drop* |

### 2.5.3. (dr)[11]

The proto-segment which has been reconstructed as *(dr) is quite a different matter, however. Judging from the scant evidence available, it seems to be regularly reflected as r in KSR:

| | | |
|---|---|---|
| *pi(dr)i | ir | *twist, twisted* |
| *si(dr)i(t) | iri | *masturbation* |

r is not a particularly frequent sound in KSR. Its retroflexed quality is unusual for Micronesia, where some sort of flap or trill is the norm. Words with r often seem to be either onomatopoeic or loanwords.

Since the regular reflex of *d appears to be KSR l, one is tempted to suggest that KSR r might reflect *r rather than *d in these two forms, at least. The fact that the *(dr)'s in these forms share a common environment /*i __ *i may or may not be relevant. More evidence is needed to decide. However, given the data at hand, one is inclined to suggest that the KSR evidence points to *r rather than *d in these forms.

### 2.5.4. *nd

*nd is the nasal-grade counterpart of *d; however, it appears to be regularly reflected in KSR as the retroflexed fricative sr.

| | | |
|---|---|---|
| *ndanu(m) | sroano- | *liquid* |
| *ndau(n) | sra | *leaf* |

A third form has been reconstructed alternately with *d and *nd:

| | | |
|---|---|---|
| *doŋo/*ndoŋo | lohng | *to hear* |

The presence of KSR l suggests that the KSR form, at least, comes from *doŋo rather than *ndoŋo.

It may seem rather odd that the nasal grade of *d should be reflected as a fricative rather than as a liquid like the oral grade reflex. However, a relationship seems to exist between liquids and affricates (rather than fricatives) in some of the Trukic languages. In Woleaian, for instance, the geminate counterpart of r is a retroflexed affricate ch [ʧ̌:] (< PMC *r), which suggests that an affricate may be the 'strengthened' counterpart of a liquid.[12] If pre-nasalisation can be regarded as a kind of 'strengthening', then it is perfectly possible that a fricative sr may have come to be the 'strengthened' counterpart of *d, since KSR lacks affricates. It is even possible that KSR did at one time possess an affricate reflex of *nd which later merged with or became sr. On present evidence, however, this is mere speculation.

### 2.5.5. *R

One might expect *R to be lost in KSR as it is in so many other OC languages. However, it is not clear that this is the case. The following is a list of all forms in the data involving *R:

| | | |
|---|---|---|
| *kuRita | koet | *octopus* |
| *waRo | ah | *rope* |
| *Ruŋma(ʔ) | yuwac-/iwac- | *house, shelter, place* |
| | lohm | *house* |
| *ʔatoluR | ahtro- | *egg* |
| *ʔaRu(s) | acsr | *current* |

In the first two forms and one (pair of) reflex(es) of the third, *R has indeed been lost. In *ʔatoluR the final consonant would probably have been lost whatever it might have been. In the form *ʔaRu(s) it is not clear whether KSR sr is supposed to reflect *R or *s. Given that POC final consonants were regularly lost in KSR (see Section 2.7.), it seems unlikely that the final KSR sr would reflect a final *s whose presence was doubtful to begin with. However, there is another form containing a non-zero reflex of *R, which is lohm < *Ruŋma(ʔ). The evidence seems a bit more clear in this second case. The fact that there are 'regular' reflexes of *Ruŋma(ʔ) (see Section 2.4.4. for reflexes of * m) suggests that lohm is an irregular reflex, possibly a borrowing (although rather an unlikely one, in view of its meaning and frequency of occurrence), possibly a case of sporadic change, or possibly dialect mixing.[13]

Whatever the case, it is clear that more could be said about the development of *R in KSR if more data were available. At the very least, one might hope to clarify the status of the various reflexes Ø, l, and sr.

## 2.6. Glides

### 2.6.1. *y

There is only one instance of *y in the data, and it has been lost in the KSR reflex of that form:

| | | |
|---|---|---|
| *mpaya | pa | *bait* |

It is not clear that *y was phonemic in POC.

### 2.6.2. *w

The reconstruction of a *w phoneme in POC appears to be more secure. *w generally is lost in KSR, although not always without leaving some trace of its presence on the vowels either preceding or following it.

| | | |
|---|---|---|
| *waka | okah | *root* |
| *nsiwa | yuh | *nine* |
| *pakiwak | pahko | *shark* |
| *waRo | ah | *rope* |
| *walu | oal- | *eight* |

In *waka > okah the sequence *wa appears to have coalesced as o. Similarly, in the word for *shark* *-iwa- appears to have become o, possibly via **ewa.[14]

Why the identical sequence in *nsiwa did not become o is not clear, although it is possible that the changes which brought about the presence of the initial glide in the KSR form yuh may have had some effect.

On the other hand, *waRo did not become o, nor did *walu become ol-. There is some disagreement among KSR speakers as to the nature of the initial vocalic portion of the word for *eight*. In some people's speech it appears to be a diphthong [ᵒɑ]. An occasional speaker might give a citation form which is best transcribed [ɒ:l], in which the symbol ɒ represents a monophthongal low back round vowel rather like that in some British speakers' pronunciation of the English word hot. In many people's speech, however, the sound spelled with the digraph oa has no rounding whatsoever; it is simply [ɑ:]. Younger speakers, in particular, appear to be losing this rounding feature.

## 2.7. Final consonants

Like other MC languages, KSR normally does not retain POC word-final consonants. For example,

| | | |
|---|---|---|
| *aŋin/*nsaŋi | eng | *wind* |
| *ʔatoluR | ahtro- | *egg* |
| *daʔa(n) | lah | *branch* |
| *ndanu(m) | sroano- | *liquid* |
| *kuli(t) | kolo-/kulu- | *skin, bark* |
| *laŋi(t) | lucng | *sky* |
| *manipi(s) | mihnini | *thin* |
| *pakiwak | pahko | *shark* |
| *tuʔu(d) | tu | *to stand* |

In transitive verbs, however, the last consonant was 'protected' from loss by the presence of a transitive suffix *-a or *-i (cf. Harrison 1978) which was subsequently lost through a general process of final vowel deletion. Thus one finds transitive-intransitive verb pairs such as

| | | | |
|---|---|---|---|
| *ŋkinit | | kihnis | *(vt) to pinch* |
| | | kihnkihn | *(vi) to pinch* |
| *mpulu(t) | | fulus | *(vt) to paste* |
| | | fulful | *(vi) to adhere, etc.* |
| *paŋ(ou)(n) | | pahngon | *(vt) to call* |
| | | pahngpahng | *(vi) to call* |
| *tanom/*tanum | | taun | *(vt) to bury* |
| | | tatuh | *(vi) to bury* |
| | | toan | *(vt) to press, bury* |
| | | toatoa | *(vi) to press, bury* |
| *tasimi | *sharpen it* | twem | *(vt) to sharpen* |
| | | twetwe | *(vi) to sharpen* |
| *ʔuluŋa | | ilung | *(vt) to rest, prop* |
| | | ilul | *(vi) to rest, prop* |

KSR kihnis is presumably descended from a proto-form *ŋkinit-i in the same fashion that KSR twem can be traced back (despite a number of phonological changes) to POC *tasimi. KSR kihnkihn and twetwe, on the other hand, may be taken to reflect *ŋkinit and *tasim, respectively, with reduplication occurring after final consonant deletion and final vowel deletion (see Section 4 for a discussion of diachronic rule ordering).

The development of ilung/ilul from *?uluŋa may represent an analogical extension of this pattern. It is possible that the final vowel on the reconstructed POC form may have been (reanalysed as) the transitive suffix *-a, in which case it would serve the same function as the final *i of *tasimi. Why the intransitive form is ilul rather than **ilil is not clear, any more than the fact that the initial vowel of both forms is i rather than u.

As for *tanom/*tanum, one would have predicted something like KSR **tanum/**tantan rather than either of the actually occurring verb pairs. The exact relation between the two pairs of KSR verbs is not clear. They may be variant pronunciations of the same forms — that is, tatuh may be merely toatoa minus rounding, as al- is oal- *eight*, minus rounding, and taun may be a variant of toan or vice versa. Both are equally 'acceptable' as reflexes of *tanom/*tanum. In both instances KSR behaves as though the last consonant of the proto-form had been n and the last syllable of the POC form had been lost at some earlier stage. The fact that MRS has jewen *press down on* suggests that the final *o/um of the POC form might not have been present at the PMC level — unless either KSR or MRS (or both) had borrowed the form from some third language. It is also possible that these particular pairs of verbs did not exist in KSR until after final consonant deletion (and probably also final vowel deletion) had applied historically, and that they were created later by a still-productive (?) process of reduplication.

Our data contain two instances of KSR transitive verbs without the final consonants reconstructed for POC:

| | | |
|---|---|---|
| *si(dr)i(t) | irii | *(vt) to masturbate* |
| *sae(t) | se | *(vt) to tear* |

In both cases the fact that the final consonant of the POC form is enclosed in parentheses indicates that it "may or may not have been part of the form" (Grace 1969:44). In neither case does the final *t appear to have been present in PMC (cf. KIR tii *to spurt*, MRS jjir *slippery, lubrication*, TRK siir *urine, to urinate* and ir *to masturbate (taboo)*, and WOL siri *to masturbate*). For *to tear* there exist alternate reconstructions without the final consonants (PEO *sai and *sei) which are included as possible Oceanic proto-forms because of the point-for-point congruence between the phonological systems reconstructed at that time for POC and PEO.

Some POC verbs ended in consonants that are normally lost in KSR:

| | | |
|---|---|---|
| *punu? | uni | *to hit, kill* |
| *?utup | uti | *(vt) to draw water* |
| | ut | *(vi) to draw water* |

The only notable feature of these forms is the fact that the final vowel is i in KSR rather than u as one might expect from a straightforward application of the process of final vowel deletion.[15]

| | |
|---|---|
| *punu? + i | ? > Ø |
| **punu + i | final vowel deletion |
| **punu | |

Evidently, something happened to the second *u. Whether it was assimilated to the following *i, which suggests that the order of the rules must be as given above with assimilation applying before final vowel deletion, or whether it was replaced by -i by analogy to the productive transitivising suffix, cannot be decided on the basis of the evidence at hand.

POC verbs which are reflected only as intransitive verbs (i.e. without transitive counterparts) show no trace of the reconstructed final consonants:

| | | |
|---|---|---|
| *matudu(R) | mutul | *to sleep* |
| *ma?udi(p) | moul | *(be) alive, life* |
| *tanji(s) | tuhng | *to cry* |

Admittedly, these are not the best possible examples, as there is some uncertainty as to the presence of the final consonant in the POC forms (although one wonders whether it is not the historical existence of a process of final consonant deletion which created this uncertainty), and there is the possibility that all of these consonants might have been lost in KSR, anyway, even had there been a following vowel to 'protect' them. Still, one would expect that KSR intransitive verbs without transitive counterparts would not retain any reconstructed POC final consonant.

## 2.8. Distinctively velarised consonants in KSR

As mentioned in Section 1., velarised consonants appear to occur non-distinctively in the environment of central and back vowels in KSR. Distinctive velarisation is restricted to syllable-initial position before a front vowel, generally the mid vowels e and ac, occasionally before i, apparently never before the low front vowel ah.

Comparison with other MC languages suggests that this distribution of velarised consonants must have developed at some post-PMC stage. Marck 1977 reconstructs a distinct set of velarised labials in PMC. KSR agrees with the other MC languages in reflecting PMC *p' (> KSR f) differently from PMC *p (> KSR p). However, PMC *m and *m' are treated identically in KSR — they are merged so that it is impossible to tell whether any given instance of synchronic m or mw comes from PMC *m or *m'. It does appear, however, that distinctively velarised consonants are especially likely to occur before KSR e < *a:

| | |
|---|---|
| *madama | mahlwem |
| *mata | muhta-/mwet |
| *tasimi | twem |

Not all instances of *a > e condition velarisation, however:

| | | | |
|---|---|---|---|
| *lima | | luhm-/lime- | *five* |
| *sae(t) | | se | *(vt) to tear* |
| *tansi(k) | *sea, salt water, salt* | te | *beach, seashore* |

The problem of the history and synchronic status of velarisation in KSR requires more consideration than can be given it here.

## 2.9. Distinctively labialised consonants in KSR

It was pointed out in Section 1., that distinctively labialised (rounded) consonants occur only before the mid front vowels e and ac. Only one form in the data contains such a sequence — KSR koet (< POC *kuRita *octopus*) — although such forms are not extraordinarily rare in KSR. No doubt a more extensive search would discover POC antecedents for a number of these words. The difficulty in this instance is not necessarily lack of data, but is due to the

unresolved question of how to analyse such sequences in the synchronic phonology. The problem of the vowel/diphthong oa may be related to this.

On the one hand, it appears as though KSR might be analysable as having no underlying diphthongs. That is, sequences such as orthographic koe and koac would be analysed as consisting of a (distinctively) rounded consonant followed by an unrounded vowel. This is the analysis implied in Lee (1975).

There is no a priori reason why a sequence such as orthographic koa could not be analysed in a similar fashion. Lee's decision to describe oa as a low back vowel (Lee 1975:15-16) may be attributed to a desire to fill an empty slot in the vowel matrix — Lee clearly describes the diphthongal pronunciation represented by this orthographic symbol.

Not all speakers actually use this diphthongal pronunciation, however. In fact, some (younger?) speakers appear to have lost all trace of rounding in words that are spelled in the dictionary with oa and also in at least some words given as having rounded consonants.[16] It may be as much a matter of theoretical prejudices as of empirical fact whether one chooses to interpret this synchronically as loss of rounding or as loss of a labial glide between the consonant and the vocalic nucleus of the syllable.

Historically, on the other hand, one would probably prefer an analysis that would account for oa as well as oe and oac. The example cited, KSR koet < POC *kuRita, does not seem to be typical of this class of sound changes. One suspects the following examples to be more representative of this kind of rounding:

| | | | |
|---|---|---|---|
| *ndanu(m) | *fresh water* | sroano- | *liquid* |
| *laŋo | | loang | *fly (insect)* |
| *napo | *surf* | noa | *wave* |
| *tanom/*tanum | | toan | *(vt) to press, bury* |
| | | toatoa | *(vi) to press, bury* |

Evidence has been given for a process of rounding assimilation in the evolution of KSR phonology (Section 3.2.3.). Excluding the operation of other processes, one would expect to find a unitary low back rounded vowel in the first syllable of these forms as a result of the application of this rule. The synchronic loss of rounding in oa suggests that this vowel is not particularly stable, perhaps because it is relatively difficult to round the lips when the mouth is open so wide. Some speakers — historically the majority? — circumvent this difficulty by diphthongising oa, pronouncing it as Lee (1975) describes. Other speakers — including many young people — find it easier to simply do away with the rounding feature. Anticipatory rounding of the preceding consonant does appear to occur when speakers pronounce oa as a round vowel or as a diphthong whose first part is rounded. When oa is unrounded, however, the preceding consonant is not rounded, either.

In the case of the mid vowels e and ac, one can imagine rounded counterparts of these vowels, probably not very stable because there would not be much support for them from the rest of the phonological system. Diphthongisation would represent a way to retain the rounding feature without adding two new vowels to the inventory.

In both cases, once diphthongisation has occurred, it would be a simple matter to transfer some features from the first part of the diphthong to the preceding consonant. There is some question as to whether the letter o in the sequences oe and oac really does represent merely a feature of the consonant

or an actual segment, a glide, occurring between the consonant and the vowel. This paper, however, is not the place to attempt to resolve the issue. What we are suggesting here is that the problem appears to exist on a diachronic as well as a synchronic level.[17]

To further complicate the issue, there are a few forms like oal (< POC *walu *eight*) in which the KSR form may or may not reflect loss of POC *w, either preceded or followed by rounding of the *a, conditioned by either the preceding *w (if rounding occurred before loss of *w) or by *u in the following syllable (presumably before final vowel deletion). Synchronically the alternation between a diphthongal pronunciation [o̯o̯al] and an unrounded pronunciation [ɑ:l] is associated with different dialect areas — Maclwem vs. the other villages. However, all KSR speakers consulted insisted that this is not a general phonological feature by which the dialect areas could be distinguished, merely a single idiosyncratic lexical item.

## 3. VOWELS

The following vowels occur in the POC reconstructions used in this paper:

| | |
|---|---|
| i | u |
| e | o |
| a | |

Their reflexes are distributed among the twelve (surface) vowels of contemporary KSR (catalogued in Section 1.) in ways which are not entirely predictable. What regularities can be found in the data are presented below.

### 3.1. Vowel correspondences

#### 3.1.1. *i

Non-final *i is most often reflected as a high vowel in KSR, particularly if the following vowel in the proto-form was also high:

| | | |
|---|---|---|
| *ŋkinit | kihnis | *(vt) to pinch* |
| *lima | luhm-/li̲me- | *five* |
| *manipi(s) | mihnini | *thin* |
| *pi(dr)i | ir | *twist* |
| *pitu | it- | *seven* |
| *pitu?u | itu/itih | *star* |
| *si(dr)i(t) | ir | *masturbation* |
| *sili | il | *to enter* |
| *anitu | inut | *ghost, spirit* |
| *niu(R) | nu | *coconut* |

Of these high vowel reflexes, KSR i is the most frequent. *i corresponds to KSR ih in only one form in the data, but this low frequency is most likely an artifact of the highly restricted database — a wider-ranging search would probably discover more instances of this historical development, which might make it possible to attempt to explain it, something one hesitates to do on the basis of a single instance.

*i appears to become u when the next vowel is *u in the forms *anitu and *niu(R), but not in *pitu or *pitu?u. It is remotely possible that the

preceding consonant may condition the application of this (diachronic) rule or possibly even a later change back from an intermediate **u to i, but there is not enough evidence here for any more precise statement.

The presence of a following *a in the proto-form seems to have caused lowering of *i in some forms:

| | | |
|---|---|---|
| *kinta | kuht | *we (1 pl)* |
| *kuRita | koet | *octopus* |
| *lima | luhm-/lime | *five* |
| *pakiwak | pahko | *shark* |
| *nsiwa | yuh | *nine* |

However, there are exceptions to this generalisation, too:

| | | |
|---|---|---|
| *kianto | kiyacs | *outrigger boom* |
| *lima | luhm-/lime- | *five* |
| *tia(n) | (in-)siyac- | *belly* |
| *tina | ninac | *mother* |

It is interesting to note that the KSR reflexes of *kianto and *tia(n) both contain a y glide between the i and the reflex of *a. When and how these glides originated is not at all clear; however, it would not be surprising to find a connection between their presence and the preservation of *i as KSR i in an environment in which one might otherwise expect lowering of *i.

This explanation will not apply to *tina and *lima (> KSR lime-), however. No non-ad hoc account of these forms is possible on the basis of the evidence at hand. There may be a way to account for the doublet from POC *lima when the KSR numeral system is more carefully examined.

In two instances medial *i appears to have been lost following *a, but not before having had some effect on that vowel:

| | | |
|---|---|---|
| *papine | acn | *woman* |
| *tasimi | twem | *(vt) to sharpen* |

The second of these forms suggests that the loss of *i in this environment must have been preceded by the loss of *s, which is conditioned by the presence of a following *i:

There is some evidence to suggest that *i may have conditioned a more general raising and sometimes fronting of a preceding *a:

| | | | |
|---|---|---|---|
| *aŋin/*nsaŋi | | eng | *wind* |
| *api | | e | *fire* |
| *kali | | kuhlkuhl | *(vi) to dig* |
| *laŋi(t) | | lucng | *sky* |
| *manipi(s) | | mihnini | *thin* |
| *nsai | | suc | *who?* |
| *tali | | sucl | *rope* |
| *taŋi(s) | | tuhng | *to cry* |
| *tansi(k) | *sea, etc.* | te | *beach, seaside* |

What determines the resulting KSR vowel is not clear, possibly a combination of the consonantal environment and whether or not the vowel is stressed. If conditioned by a following *i, the raising rule must have applied before final vowel deletion removed the conditioning factor in many of the forms above.

The reciprocal prefix *pai- is the sole exception (in the data) to this raising rule, as its KSR reflex is a-.

*i evidently had no effect on preceding non-low vowels:

| | | |
|---|---|---|
| *mpoŋi | fong | *night* |
| *pu(n)ti | usr | *banana* |
| *sili | il | *to enter* |
| *tuki | tuk | *(vt) to pound* |

In only one instance is non-final *i lost without any trace:

| | | |
|---|---|---|
| *inu(m) | nihm | *(vt) to drink* |

Evidence from other MC languages suggests that this development must have occurred at some proto- or pre-MC stage (cf. KIR nina *(vt) drink, swallow*, MRS limiy *(vt) to drink*, PNP nim *(vt) drink*, ULI lema *(poss cl) drink*).

### 3.1.2. *e

*e is indisputably the lowest-frequency vowel in the POC reconstructions on which this paper is based. Only one form had a KSR reflex of non-final *e:

| | | |
|---|---|---|
| *mpeŋka | fak | *bat, flying fox* |

A handful of forms suggest that *e might have had an effect on preceding *a similar to that of *i:

| | | |
|---|---|---|
| *?ate | acsyac-/esyac- | *liver* |
| *(n)sake | -yak/-ack | *up, upwards* |
| *sae(t) | se | *(vt) to tear* |
| *mpule | ful | *shell: cowry* |
| *papine | acn | *woman* |

### 3.1.3. *u

Like *i, *u is most often reflected as a high vowel in KSR, most frequently u:

| | | |
|---|---|---|
| *kutu | kut | *louse* |
| *mpulu(t) | fulus | *(vt) to paste, gum* |
| *pu(n)ti | usr | *banana* |
| *sulu | sul | *torch* |
| *tu?u(d) | tu | *to stand* |
| *?umu | um | *earth oven* |
| *?una(p) | unac-/une- | *fish scale, etc.* |

but sometimes i or ih:

| | | |
|---|---|---|
| *mputo | fuht/fihtac- | *navel* |
| *Ruŋma(?) | yuwac-/iwac- | *house, etc.* |
| *susu | titi- | *breast* |
| *?uluŋa | ilung | *(to) pillow* |
| *?uma | imac | *garden, etc.* |
| *inu(m) | nihm | *(vt) to drink* |

Again like *i, *u is sometimes reflected by a lower vowel in KSR, apparently due to the presence of a low vowel in the following syllable:

| | | |
|---|---|---|
| *dua | lo/luo/lu- | *two* |
| *muta(?) | woht[18] | *(vi) to vomit* |
| *Ruŋma(?) | lohm | *house* |

There are exceptions to this, however:

| | | |
|---|---|---|
| *dua | lo/luo/lu- | *two* |
| *?una(p) | unac-/une- | *fish scale, etc.* |
| *?uma | imac | *garden, etc.* |
| *muta(?) | wihte[19] | *(vt) to vomit* |

as well as one and a half instances of unexplained lowering:

| | | |
|---|---|---|
| *?atolu(R) | ahtro- | *egg* |
| *paŋ(ou)(n)[20] | pahngon | *(vt) to call* |

A somewhat better case can be made for *u having conditioned raising and/or rounding of the vowel in the preceding syllable:

| | | |
|---|---|---|
| *ndanu(m) | sroano- | *liquid* |
| *manu(k) | won | *bird* |
| *matudu(R) | mutul | *to sleep* |
| *ma?udi(p) | moul | *alive, life* |
| *ŋapulu(?) | (si-)nguhul | *ten* |
| *anitu | inut | *ghost, spirit* |
| *niu(R) | nu | *coconut* |
| *patu | yot | *stone* |
| *taku | tok/tohkoh/tuku- | *back (anat.)* |

However, there are exceptions to this, too:

| | | |
|---|---|---|
| *?anusi | acni | *to spit* |
| *?aRu(s) | acsr | *current, stream* |
| *?atolu(R) | ahtro- | *egg* |
| *ndau(n) | sra | *leaf* |
| *natu | nahtuh-/nuhtV- | *child* |
| *panua | acn | *land* |

for which no systematic explanation is currently available.

### 3.1.4. *o

Non-final POC *o is generally reflected as o in KSR, occasionally as oh, occasionally as a non-round mid vowel:

| | | |
|---|---|---|
| *matolu | mahtol | *thick* |
| *onom | on-/ohn- | *six* |
| *paŋ(ou)(n) | pahngon | *(vt) to call* |
| *mpo- | fo | *smell* |
| *mponot | fonos | *(vt) to block, etc.* |
| *mpoŋi | fong | *night* |
| *mpou | fo | *pole, etc.* |
| *tolu | tol- | *three* |
| *doŋo | lohng | *to hear* |
| *topu | tuh | *sugarcane* |
| *?ulo(s) | ulac | *maggot* |

A possible explanation for the unrounding of the vowel in tuh (< *topu, which contains two round vowels) may be found in the fact that the word tohoh also exists in KSR as the name of a kind of tree. This may be an instance of avoidance of ambiguity.

The presence of a final vowel in ulac is unexpected to begin with. Ordinarily the last vowel would have been lost, whether or not the final *s was actually present, because final vowel deletion seems to have followed final consonant deletion historically (see Section 4. for further discussion of diachronic rule ordering). In a so-called 'alienable' noun like ulac (see Section 3.3.3. for further discussion of 'alienable' vs. 'inalienable' nouns), there would have been no possessive suffix to 'protect' the historical final vowel from deletion. This suggests that the final vowel in KSR ulac does not reflect the last vowel of POC *ʔulo(s), but there is no evidence to indicate where it might have come from.

The evidence concerning the effects of a following *o on a preceding non-round vowel is contradictory. In some instances it appears to have no effect:

| | | | |
|---|---|---|---|
| *aŋoaŋo | *yellow* | ahng | *turmeric* |
| *kianto | | kiyacs | *outrigger boom* |
| *mapo | | mah | *to heal* |
| *tanom/*tanum | | taun | *(vt) to bury* |
| | | tatuh | *(vi) to bury* |
| *waRo | | ah | *string, rope* |

Sometimes, however, it appears to result in rounding of the vowel in the preceding syllable, in forms very much like those listed above:

| | | | |
|---|---|---|---|
| *laŋo | | loang | *fly (insect)* |
| *napo | *surf* | noa | *wave* |
| *tanom/*tanum | | toan | *(vt) to press, bury* |
| | | toatoa | *(vi) to press, bury* |

There is always the possibility that the surrounding consonants may have an effect on the vowel, although this remains to be demonstrated. There is also the possibility that these two different sets of forms represent different historical strains in the development of KSR. As mentioned in Section 2.7., the relationship between taun/tatuh and toan/toatoa has not yet been explained satisfactorily — it may be that different dialects of KSR varied in their application of the rounding assimilation rule and that the present situation is a result of dialect mixing.

In all of the examples above the vowel which could have been influenced by *o was *a. That is not because there is any a priori reason to expect that only *a was susceptible to such influence, but because the data contain no words of the shape *CiCo or *CeCo. Obviously, words of the shape *CoCo or *CuCo would not constitute evidence either for or against a historical rule of rounding assimilation unless that rule were supposed to have operated at a Pre-POC stage which is well beyond the domain of this paper.

### 3.1.5. *a

*a may be reflected as any of the twelve surface vowels in KSR. Generally, it is a non-high, but not necessarily low, vowel. More often than not, it is unrounded.

**low vowel reflexes**

*a > ah

| | | | |
|---|---|---|---|
| *aŋoaŋo | *yellow* | ahng | *turmeric* |
| *?atoluR | | ahtro- | *egg* |
| *da?a(n) | | lah | *branch* |
| *madama | | mahlwem | *moon* |
| *mapo | | mah | *to heal* |
| *natu | | nahtuh-/nuhtV- | *child* |
| *paka- | | ahk- | causative pref |
| *waka > PMC *wakaRa | | okah | *root* |

*a > a

| | | | |
|---|---|---|---|
| *ndau(n) | | sra | *leaf* |
| *ŋmata | | wet/wat | *worm* |
| *pai- | | a- | reciprocal pref |
| *mpaya | | pa | *bait* |
| *nsakaRu | *reef* | tuhka | *island* |
| *(n)sake | | -yak/-ack | *up, upwards* |
| *tama | | tuhma- | *father* |
| *tanom/*tanum | | taun | *(vt) to bury* |
| | | tatuh | *(vi) to bury* |

*a > oa

| | | | |
|---|---|---|---|
| *ndanu(m) | | sroano- | *liquid* |
| *laŋo | | loang | *fly (insect)* |
| *napo | *surf* | noa | *wave* |
| *tanom/ tanum | | toan | *(vt) to press, bury* |
| | | toatoa | *(vi) to press, bury* |
| *walu | | oal- | *eight* |

Notice that in all cases where *a > oa the vowel in the following syllable is round. The converse is not true, however; that is, a round vowel in the second or subsequent syllable of a proto-form does not necessarily cause rounding of *a in the preceding syllable. No ready explanation offers itself for the choice between a and ah reflexes of *a.

**low-mid vowel reflexes**

*a > ac

| | | |
|---|---|---|
| *?anusi | acni | *to spit* |
| *?aRu(s) | acsr | *current, stream* |
| *panua | acn | *land, earth* |
| *papine | acn | *woman* |
| *tia(n) | (in-)siyac- | *belly* |
| *?una(p) | unac-/une- | *body hair, etc.* |
| *(n)sake | -yak/-ack | *up, upwards* |
| *kianto | kiyacs | *outrigger boom* |
| *Ruŋma(?) | yuwac-/iwac- | *house, shelter* |
| *?ate | acsyac-/esyac- | *liver* |
| *tina | ninac | *mother* |
| *?uma | imac | *garden, etc.* |

*a > uh

| | | |
|---|---|---|
| *mata | muht̲a-/mwet | *eye, face* |
| *natu | nah̲tuh-/nuht̲V- | *child* |
| *nsakaRu | tuh̲ka | *island* |
| *tama | tuh̲ma- | *father* |
| *taŋi(s) | tuh̲ng | *to cry* |

*a > oh

| | | |
|---|---|---|
| *taku | tok/toh̲koh-/tuku- | *back (anat.)* |

Among the lower mid vowel reflexes of *a the front vowel seems to predominate. As before, the only instance in which *a is reflected as a round vowel is when the vowel in the following syllable of the proto-form is round. Contrary to the case of the low vowel reflexes of *a, there appears to be a synchronic explanation for the choice between one lower mid reflex and another. Except for tuhng, all of the words with KSR uh as a reflex of *a have that vowel in unstressed position (see Lee 1975:33-35 for a summary of KSR stress patterns). The fact that the KSR word for *to cry* is tuhng rather than tucng (see examples below) remains inexplicable at present.

high-mid vowel reflexes

*a > e

| | | |
|---|---|---|
| *aŋin/ nsaŋi | eng | *wind* |
| *api | e | *fire* |
| *?ate | acs̲yac-/e̲syac- | *liver* |
| *madama | mahlwe̲m | *moon* |
| *mata | muhta-/mwe̲t | *eye, face* |
| *ŋmata | we̲t/wat | *worm* |
| *sae(t)/*sai/*sei | se | *(vt) to tear* |
| *tasimi | twem | *(vt) to sharpen* |
| *tansi(k) *sea, etc.* | te | *beach, seaside* |
| *?una(p) | unac-/une- | *body hair, etc.* |

*a > uc

| | | |
|---|---|---|
| *kali | kuhlkuhl [kə:lkə:l][21] | *(vi) to dig* |
| *laŋi(t) | lucng | *sky* |
| *ŋapulu(?) | singuhul [si:ŋə:u:l] | *ten* |
| *tali | sucl | *rope* |
| *nsai | suc | *who?* |

*a > o

| | | |
|---|---|---|
| *taku | to̲k/tohkoh-/tuku- | *back (anat.)* |
| *manu(k) | wo̲n | *bird* |
| *ma?udi(p) | moul | *life, alive* |
| *patu | yot | *stone* |
| *waka | o̲kah | *root* |

As before, rounded reflexes of *a occur in words with a round vowel in the following syllable, except for okah, in which the o may be taken to result from the coalescence of the sequence *wa.

In addition to this, there appears to be some evidence here for a historical process of low-vowel dissimilation:

*a > e / _____*Ca

as exemplified by the KSR forms mahlwem, mwet, and wet. The non-occurrence of forms like **melwem suggests that the rule must have applied from right to left. The lower mid front vowel (orthographic ac) never seems to be the result of low-vowel dissimilation, nor do the central or back mid vowels.[22]

Most often, KSR e, uc, and o reflexes of POC *a seem to have resulted from an assimilative process of *a-raising conditioned by the presence of a high vowel in the following syllable. If that following vowel is round, the result is usually a round vowel (although not in the case of singuhul < *ŋapulu(?)). e frequently alternates with ac as the stem vowel of the so-called 'inalienable' nouns, apparently in free variation, based on the (synchronic) evidence to date.

high vowel reflexes

| | | | |
|---|---|---|---|
| *a > i | | | |
| | *anitu | inut | *ghost, spirit* |
| *a > ih | | | |
| | *manipi(s) | mihnini | *thin* |
| *a > u | | | |
| | *matudu(R) | mutul | *to sleep* |
| | *taku | tok/tohtoh-/tuku- | *back (anat.)* |

One would not normally expect to find a proto-low vowel reflected as a high vowel in a daughter language (although the Great English Vowel Shift demonstrates that stranger things have happened). At any rate, it will not surprise most readers that KSR high vowel reflexes of POC *a are so few. What seems to be significant about these reflexes is the fact that they all occur in the unstressed (first) syllable of words whose stressed (second) syllable contains a (synchronic as well as historical) high vowel. This suggests that there might have been a second (whether synchronic or historical is not clear) assimilative vowel-raising rule which operated on short unstressed vowels (most likely mid vowels resulting from the other — presumably earlier — rule), causing them to become high vowels when followed by a stressed high vowel. Considerably more data would be needed to test this hypothesis.

## 3.2. Vowel assimilation rules[23]

The vowel correspondences catalogued above seem to suggest that a number of vowel assimilation rules must have applied in the development of KSR from POC. Most prominent among these are vowel height assimilation (both raising of low vowels and lowering of high vowels) and rounding assimilation. There also seems to be some evidence for a process of monophthongisation.

Other developments involving the vowels are low-vowel dissimilation (see Section 3.3.1.), glide insertion/glide formation (3.3.2.), and final vowel deletion (3.3.3.).

### 3.2.1. *a-raising

Evidence has already been presented in Section 3.1.5. which indicates that there might have been at least two rules having the effect of causing POC *a

to be reflected as a higher vowel in KSR. One of these rules (historically the earlier of the two?) had the effect of raising *a to a mid vowel (ac, uh, oh, e, uc, o) when the vowel in the following syllable was a high vowel. Aside from rounding assimilation, which will be discussed more fully in Section 3.2.3., it is not clear what determines which of the mid vowels would develop in any particular instance. Examples of this first raising rule are repeated here for the convenience of the reader:

| | | |
|---|---|---|
| *?anusi | acni | *to spit* |
| *papine | acn | *woman* |
| *taŋi(s) | tuhng | *to cry* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *aŋin/*nsaŋi | eng | *wind* |
| *tasimi | twem | *(vt) to sharpen* |
| *laŋi(t) | lucng | *sky* |
| *nsai | suc | *who?* |
| *manu(k) | won | *bird* |
| *patu | yot | *stone* |

The second of the raising rules appears to have applied only to unstressed vowels in the first syllable of a word — which need not necessarily have been low vowels at the time that the rule applied, although they seem to have come originally from *a — followed by a stressed high vowel in the next syllable. The output of the rule was an unstressed short high vowel (examples repeated from Section 3.1.):

| | | |
|---|---|---|
| *anitu | inut | *ghost, spirit* |
| *manipi(s) | mihnini | *thin* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *matudu(R) | mutul | *to sleep* |

At least the first of these rules is not without exceptions:

| | | |
|---|---|---|
| *ndau(n) | sra | *leaf* |
| *pai- | a- | reciprocal pref |
| *nsakaRu *reef* | tuhka | *island* |
| *natu | nahtuh-/nuhtV- | *child* |
| *ndanu(m) | sroano- | *liquid* |
| *walu | oal- | *eight* |

However, no systematic explanation for these is immediately apparent.

### 3.2.2. High vowel lowering

The evidence for a rule lowering POC high vowels to KSR mid vowels when followed by *a in the next syllable is rather less satisfactory than that given above for low-vowel raising. Favouring the existence of such a rule are forms such as

| | | |
|---|---|---|
| *kinta | kuht | *we (incl)* |
| *kuRita | koet | *octopus* |
| *lima | luhm/lime- | *five* |
| *pakiwak | pahko | *shark* |
| *nsiwa | yuh | *nine* |
| *dua | lo/luo/lu- | *two* |
| *muta(?) | woht | *(vi) to vomit* |
| *Ruŋma(?) | lohm | *house* |

Almost as many forms can be found that would have to be regarded as exceptions or counterexamples to this rule:

| | | |
|---|---|---|
| *kianto | kiyacs | *outrigger boom* |
| *tia(n) | (in-)siyac- | *belly* |
| *tina | ninac | *mother* |
| *lima | luhm/lime- | *five* |
| *dua | lo/luo/lu- | *two* |
| *ʔuna(p) | unac-/une- | *fish scales, etc.* |
| *ʔuma | imac | *garden, etc.* |

Again, no systematic explanation immediately suggests itself. It is quite possible that there is some other way to account for this distribution of reflexes of the POC high vowels which might be uncovered by further investigation.

### 3.2.3. Rounding assimilation

On the other hand, the evidence is relatively good for a historical process of rounding assimilation whereby unrounded proto-vowels became rounded when the vowel in the following syllable was round:

| | | |
|---|---|---|
| *ndanu(m) | sroano- | *liquid* |
| *manu(k) | won | *bird* |
| *matudu(R) | mutul | *to sleep* |
| *maʔudi(p) | moul | *life, alive* |
| *anitu | inut | *ghost, spirit* |
| *niu(R) | nu | *coconut* |
| *patu | yot | *stone* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *laŋo | loang | *fly (insect)* |
| *napo *surf* | noa | *wave* |
| *tanom/*tanum | toan | *(vt) to press, bury* |
| | toatoa | *(vi) to press, bury* |

But one can find just as many exceptions to this generalisation:

| | | |
|---|---|---|
| *ʔanusi | acni | *to spit* |
| *ʔaRu(s) | acsr | *current, stream* |
| *ʔatolu(R) | ahtro- | *egg* |
| *ndau(n) | sra | *leaf* |
| *natu | nahtuh-/nuhtV- | *child* |
| *panua | acn | *land, earth* |
| *aŋoaŋo | ahng | *turmeric* |
| *kianto | kiyacs | *outrigger boom* |
| *mapo | mah | *to heal* |
| *tanom/*tanum | taun | *(vt) to bury* |
| | tatuh | *(vi) to bury* |
| *waRo | ah | *string, rope* |

It would be interesting to see if the exceptions to one rule also turned out to be exceptions to other rules; however, this possibility cannot be explored in this paper, owing to limitations on time and data.

### 3.2.4. Monophthongisation

The evidence is scanty, but a few forms in the data suggest that some kind of monophthongisation or coalescence rule may have applied in the history of KSR. The forms that most strongly suggest the existence of such a rule are

| | | |
|---|---|---|
| *pakiwak | pahko | *shark* |
| *papine | acn | *woman* |
| *tasimi | twem | *(vt) to sharpen* |
| *waka | okah | *root* |

From the first three examples it appears that either the second vowel (which became juxtaposed to the first as a result of the loss of the intervening proto-consonant) in a VV sequence, or the glide in the sequence VG, first conditioned assimilation of the first vowel (for height and/or roundness) and then was lost through either deletion or coalescence — e.g. either

| | |
|---|---|
| *papine | *p > Ø |
| **aine | final vowel deletion |
| **ain | *a-raising |
| **acin | V > Ø / V — |
| acn | monosyllabic lengthening[24] |
| [ɛ:n] | |

or

| | |
|---|---|
| *papine | *p > Ø |
| **aine | final vowel deletion |
| **ain | monophthongisation |
| acn [ɛ:n] | |

For reasons of economy at least, one would probably want to opt for the second of these derivations, but it is not clear from the data that things were necessarily so.

The fourth form cited above differs from the first three primarily in that, in this case, the glide precedes the vowel and the resulting vowel may be short rather than long (although the question of vowel length in KSR has yet to be resolved on either a historical or a synchronic basis). Perhaps the rule ought to be written to accommodate both orders V-Vc and Vc-V (where the symbol Vc stands for the segment — vowel or glide — that seems to 'disappear' from the proto-form).

Again, there are exceptions, forms such as

| | | |
|---|---|---|
| *ŋmata | wet/wat | *worm* |
| *tanom/ tanum | taun | *(vt) to bury* |

as well as synchronic forms like paing *to greet*, sauk *to catch, grasp*, wal *proper, suitable*, vat *eastern part of a village*, etc. which suggest that this rule of coalescence was probably rather restricted in its application and, furthermore, is no longer active.

## 3.3. Other historical developments involving POC vowels

### 3.3.1. Low vowel dissimilation

The evidence for such a rule is scanty in KSR (although it appears to form part of the history of other nuclear Micronesian languages).

| | | |
|---|---|---|
| *madama | mahlwem | *moon* |
| *mata | muhta-/mwet | *eye, face* |
| *ŋmata | wet/wat | *worm* |

Evidence against a rule of low vowel dissimilation is provided by forms such as

| | | |
|---|---|---|
| *da?a(n) | lah | *branch* |
| *paka- | ahk- | causative pref |
| *mpaya | pa | *bait* |
| *nsakaRu *reef* | tuhka[25] | *island* |

A more extensive database is clearly necessary for the resolution of this question.

### 3.3.2. Glide insertion/glide formation

The question of the status of glides in POC has not yet been resolved to everyone's satisfaction, nor has it been entirely resolved with respect to the synchronic phonology of KSR — all of which means that any statement about developments involving glides must necessarily be highly tentative. Nonetheless, it seems clear enough that the glides are distributed differently in present-day KSR than in the ancestral language. The reconstructed POC glides all have been lost (see Section 2.6.), while other glides have arisen in different places. Not all of the glides in present-day KSR can be accounted for on the basis of the few forms considered in this paper, but some guesses can be hazarded.

Some instances of KSR w appear to be merely irregular reflexes of POC *m or (regular reflexes of) *ŋm:

| | | |
|---|---|---|
| *manu(k) | won | *bird* |
| *ŋmata | wet/wat | *worm* |
| *muta(?) | wihte | *(vt) to vomit* |
| | woht | *(vi) to vomit* |
| *Ruŋma(?) | yuwac-/iwac- | *house, shelter* |

Some instances of KSR y appear to be successors of lost proto-consonants:

| | | |
|---|---|---|
| *patu | yot | *stone* |
| *Ruŋma(?) | yuwac-/iwac- | *house, shelter* |
| *(n)sake | -yak/-ack | *up, upwards* |
| *nsiwa | yuh | *nine* |

although the exact mechanism by which these arise remains unclear.

Still other instances of KSR y appear to be epenthetic.

Forms such as

| | | |
|---|---|---|
| *kianto | kiyacs | *outrigger boom* |
| *tia(n) | (in-)siyac- | *belly* |

suggest that there may have been a rule in KSR inserting a y-glide between a high front vowel and a following (lower) mid front vowel or possibly a historical rule inserting y between a historical high front vowel and a historical low vowel. The data at hand do not even permit a statement of whether or not a similar rule applied with respect to historical sequences of high (rounded) back vowel followed by low vowel, much less any statement concerning the synchronic as opposed to diachronic status of such a rule.

Further complications arise from the as yet unresolved nature of such features of synchronic KSR phonology as the rounding on the consonant or the first part of the diphthong (if such it be) in forms like the following:[26]

| | | |
|---|---|---|
| koet | < *kuRita | *octopus* |
| sroano- | < *ndanu(m) | *liquid* |
| loang | < *laŋo | *fly (insect)* |
| noa | < *napo | *wave (< surf)* |

The problem of glides clearly warrants further investigation.

### 3.3.3. Final vowel deletion

In addition to final consonant deletion (see Section 2.7.), there appears to have been a rule in the development of KSR deleting final vowels from POC forms. This rule evidently applied after final consonant deletion (see Section 4 for discussion of diachronic rule-ordering), resulting in many consonant-final words in KSR:

| | | |
|---|---|---|
| *anitu | inut | *ghost, spirit* |
| *aŋin/*nsaŋi | eng | *wind* |
| *ndau(n) > **ndau > **nda > sra[27] | | *leaf* |
| *kinta | kuht | *we (incl)* |
| *kuRita > **kuita > **kueta > **kuet > koet | | *octopus* |
| *pakiwak > **pakiwa > **pakewa > **pakew > pahko | | *shark* |
| *patu | yot | *stone* |
| *doŋo | lohng | *to hear* |
| *matolu | mahtol | *thick* |
| *pi(dr)i | ir | *twist(ed)* |
| *sili | il | *to enter* |
| *kali > **kuhli > **kuhl > kuhlkuhl | | *(vi) to dig* |
| *tasimi | twem | *(vt) to sharpen* |

As indicated by the above examples, this process appears to have applied to words of all classes impartially.

However, as with final consonant deletion, there appears to be an entire class of cases in which the final vowel of the POC form has been 'protected' by the presence of a suffix. This class consists of 'inalienable' nouns, those nouns for which possession is indicated by the direct attachment of a possessive suffix (as opposed to the 'alienable' nouns, for which possession is indicated by a following 'classifier' carrying the possessive suffix). The vowel that was affected by final vowel deletion was the final vowel of the possessive suffix rather than that of the noun stem. Thus we find the final vowel of the reconstructed proto-form reflected as the stem(-final) vowel of the inalienable noun:

| | | |
|---|---|---|
| *tama | tuhma- | *father* |
| *mata | muhta-/mwet | *eye, face* |
| *natu | nahtuh-/nuhtV- | *child* |
| *mputo | fihtac-/fuht | *navel* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *ʔate | acsyac-/esyac- | *liver* |

The alternation in the vowels of the contemporary KSR forms may be attributed (at least historically) to the influence of the vowel in the possessive suffix.

## 4. DIACHRONIC RULE-ORDERING

In preceding sections of this paper we have tentatively put forward a number of rules that might have applied during the development of the KSR phonological system. It seems quite clear that the changes represented by these rules did not all occur at one and the same time. In this section we will examine the data for evidence bearing on historical rule ordering.

Final consonant deletion may have occurred quite early in the history of KSR. It must have preceded final vowel deletion, if one can assume the final consonant to have been present in the following POC forms:

| | | |
|---|---|---|
| *aŋin/*nsaŋi | eng | *wind* |
| *laŋi(t) | lucng | *sky, heaven* |
| *manu(k) | won | *bird* |
| *matudu(R) | mutul | *to sleep* |
| *ma?udi(p) | moul | *alive, to live* |
| *taŋi(s) | tuhng | *to cry* |
| *tudu(?) | tu | *to stand* |

This ordering is supported by the existence of forms in which a final reconstructed vowel was lost but the consonant that preceded it has been retained, e.g.:

| | | |
|---|---|---|
| *anitu | inut | *ghost, spirit* |
| *doŋo/*ndoŋo | lohng | *to hear* |
| *kianto | kiyacs | *outrigger boom* |
| *madama | mahlwem | *moon* |
| *mpeŋka | fak | *bat, flying fox* |
| *tali | sucl | *rope* |

A number of forms seem to suggest that final consonant deletion, but not final vowel deletion, must have taken place before the development of the inalienable possessive marking system. For example,

| | | |
|---|---|---|
| *ndanu(m) | sroano- | *liquid* |
| *kuli(t) | kolo-/kulu- | *skin, bark* |
| *Ruŋma(?) | yuwac-/iwac- | *house, shelter* |
| *tia(n) | (in-)siyac- | *belly* |
| *?una(p) | unac-/une- | *fish scale, etc.* |

The presence of the possessive suffix would 'protect' the final vowel of vowel-final forms from deletion:

| | | |
|---|---|---|
| *?ate | acsyac-/esyac- | *liver* |
| *mata | muhta-/mwet | *eye, face* |
| *natu | nahtuh-/nuhtV- | *child* |
| *mputo | fuht/fihtac- | *navel* |
| *susu | titi- | *breast* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *tama | tuhma- | *father* |

These forms support the hypothesis that the development of inalienable possession marking preceded final vowel deletion.

A number of changes are conditioned by a following vowel. There is evidence that these, too, must have occurred before final vowel deletion. For instance, in forms such as

| | | |
|---|---|---|
| *ŋkinit | kihnis | *(vt) to pinch* |
| *mponot/*ponot | fonos | *(vt) to block, plug* |
| *mpulu(t) | fulus | *(vt) to paste* |

the final -s in the KSR forms presumably reflects POC *t, which became s in the presence of a following *-i transitive suffix (see Section 2.1.2.), which was later deleted along with other final vowels.

The set of transitive verbs cited above also provides evidence that the *t > s change must have followed the change of *s > Ø / ___ *i (see Section 2.3.1.) — or else that the latter change did not occur until after final vowel deletion had taken place. To suppose that a final *i conditioned first a change of *t to s and then *s > Ø before being deleted would leave the final s in these forms unaccounted for. However, the data used for this paper do not permit us to clarify the matter. Only one item in the data is reconstructed as ending in *-si, POC *ʔanusi *to spit*. This is reflected in KSR as acni. The absence of s suggests that *s did become Ø. Since we have established that final consonant deletion preceded final vowel deletion (and if we can take the synchronic final -i to be an irregular reflex of *u rather than a reflex of the POC final vowel), the only rule that could have deleted *s presumably was the one conditioned by following *i. Thus we argue that the loss of *s / ___ *i must have antedated not only final vowel deletion but also the change of *t > s.

The loss of *s /____*i also must have occurred before monophthongisation (see Section 3.2.4.), as evidenced by KSR twem *(vt) to sew* (< POC *tasimi):

| | |
|---|---|
| *tasimi | *s > Ø / __ *i |
| **taimi | final vowel deletion |
| **taim | monophthongisation/velarisation of the preceding consonant |
| twem | |

Loss of *ns and *p also may have preceded monophthongisation. For example,

| | |
|---|---|
| *tansi(k) | final consonant deletion |
| **tansi | *ns > Ø |
| **tai | monophthongisation |
| te[28] | |

(See also the derivation given for acn *woman* < POC *papine in Section 3.2.4.). KSR e *fire* may have evolved from POC *api in a similar fashion.[29]

Among the other changes that must have antedated final vowel deletion are *a-raising, high-vowel lowering, rounding assimilation, and low-vowel dissimilation. The examples below give evidence that all of these rules may have been conditioned by a final vowel which was subsequently lost.

*a-raising

| | | |
|---|---|---|
| *aŋin/*nsaŋi | eng | *wind* |
| *laŋi(t) | lucng | *sky, heaven* |
| *tali | sucl | *rope* |
| *taŋi(s) | tuhng | *to cry* |

high-vowel lowering

| | | |
|---|---|---|
| *kinta | kuht | *we (1 pl)* |
| *kuRita | koet | *octopus* |
| *muta(?) | woht | *(vi) to vomit* |
| *Ruŋma(?) | lohm[30] | *house* |
| *nsiwa | yuh | *nine* |

rounding assimilation

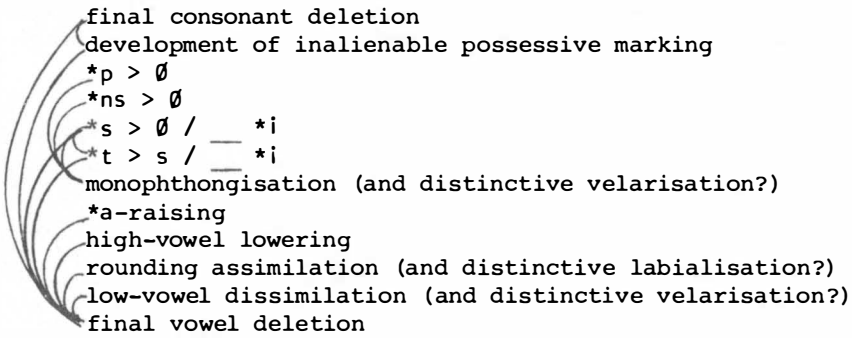
| | | |
|---|---|---|
| *laŋo | loang | *fly (insect)* |
| *manu(k) | won | *bird* |
| *napo *surf* | noa | *wave* |
| *taku | tok/tohkoh-/tuku- | *back (anat.)* |
| *tanom/*tanum | toan | *(vt) to press, bury* |

low-vowel dissimilation

| | | |
|---|---|---|
| *madama | mahlwem | *moon* |
| *mata | muhta-/mwet | *eye, face* |
| *ŋmata | wet/wat | *worm* |

No evidence can be found in the data to order the various other consonantal changes with respect to any of the rules described in this paper. This includes the loss of *? and *R, and the various replacements (*d > l, *nd > sr, etc.). The development of distinctive velarisation and labialisation may be connected with low-vowel dissimilation (and monophthongisation?) and rounding assimilation, respectively. The development of glides in KSR warrants another paper in itself.

The following schema sums up what we have been able to observe about diachronic rule ordering in KSR (following the old transformational-generative convention, pairs of rules that can be ordered with respect to each other are connected by lines):




unordered with respect to the above:

other consonantal changes
(*? > Ø, *R > Ø, *d > l, *nd > sr, etc.)
development of glides

late or synchronic:[31]

reduplication
-i transitive derivation
etc.

## 5. UNANSWERED QUESTIONS

It is difficult to come to any conclusions in a paper of this sort. No general questions have been asked or answered, only the specifics of how the various phonemes that have been reconstructed for POC appear to be reflected in KSR. With a sample as small as that used for this paper, it is not easy (or

perhaps all too easy) to generalise from the data. Clearly more work needs to be done, both synchronic and diachronic, before we will know enough about the evolution of KSR phonology to be able to attempt a clarification of relationships between KSR and other MC languages.

Many of the unanswered — perhaps unanswerable — questions about the history of KSR concern the amount of variation in the phonology and the multiple reflexes of the various POC phonemes. We have speculated about the possibility of dialect mixing and/or borrowing from as yet unidentified sources. These are among the areas in which work needs to be done.

## POSTSCRIPT

Since the time when Kee-dong Lee wrote the original version of this paper, much work has been done in both synchronic and diachronic MC and OC linguistics. I have access to the KSR dictionary and reference grammar which Lee did not complete until well after he wrote his term paper for Professor Grace's course. The one thing that still is lacking is a good description of the synchronic phonology of KSR. Many of the questions raised but not answered in this paper will not be resolved without an understanding of the underlying representations of the synchronic forms involved. The orthography used in the dictionary is adequate for most purposes, but it is not phonemic.

I have been attempting to attack the KSR problem from both sides, synchronic and diachronic. The small samples of KSR speakers with whom Lee and I have worked do not allow us to solve the problem of synchronic variation. We recognise that variation exists and suspect that it may lie behind some of the complications in the phonological system. Perhaps if it were possible to sort out the variation, it might be possible to establish a 'core' phonological system shared by all KSR speakers, from which individual speakers may deviate in idiosyncratic (?) ways. No one has ever found much evidence for local dialect variation in present-day KSR, although such variation probably existed in pre-European contact times, as in most other Micronesian speech communities. The historical records are poor. It most likely never will be known exactly what happened to the population of Kosrae as it dwindled from early estimates of 2000 - 5000 to less than 200 native Kosraeans in 1880. A recent count gives a population of 4900, but does not specify how many of these are native Kosraeans. However many native Kosraeans there may be at present, they are almost certainly not distributed according to the original precontact dialect groups. In the absence of extensive written records, the only information available is whatever might have been passed via oral tradition — including, of course, the language itself. To my knowledge no one has attempted any thorough study of the oral traditions of Kosrae. It will soon be too late — if it is not already too late — to do so. The younger people are not learning either the old ways or the old language. Both the society and its language have been changing rapidly. The lexicon is the most obvious aspect of this linguistic change, but informants note differences in pronunciation between older and younger speakers (cf. Lee 1975:391-392). No observations have been made of differences in syntax, but that is not to say that such differences may not come to light.

There is another possible explanation for some of the variation in KSR — that is, that present-day KSR may reflect a mixed linguistic heritage, not merely in terms of dialect mixing, but also mixing in elements from some other language (or languages?). This is not necessarily to say that KSR is a 'mixed language' — any more than is a language such as Rotuman (Biggs 1965).

I have begun re-examining KSR reflexes of POC phonemes using a larger sample than was available to Kee-dong Lee in 1972. The reconstructed POC lexicon is considerably larger than it was then, and by being more flexible about cognacy — at least as a first step — than Lee allowed himself to be, I am finding that KSR has multiple reflexes of many POC phonemes. Of these, there often is one that might be considered the 'regular' reflex. It seems to be the most frequent and generally agrees with the reflexes in the other MC languages. Of the minor reflexes, a few can be explained away as sporadic changes, but there are others that occur too frequently to be so lightly dismissed. A preliminary examination of these minority reflexes suggests that there may be a tendency for them to co-occur with each other in given lexical items, and not with the majority set. The next step is to take the entire corpus of POC reconstructions available to me and examine the KSR reflexes of that set. If my preliminary findings are borne out by the larger set, KSR will join Rotuman as another definitely identified case of direct and indirect inheritance in OC. Unlike the Rotuman situation, however, the source of the indirect inheritance in KSR is not particularly evident. The suspect forms do not immediately call to mind any other MC language. Nor do they look especially Polynesian, although Nukuoro and Kapingamarangi — and Kiribati — testify to the presence of Polynesians in Micronesia. Nauruan is a possibility which needs to be examined more closely. The next most likely candidates lie in Melanesia, but not much is known about many of these languages.

In terms of the nuclear Micronesian languages as a group,[32] sorting out the different sets of KSR reflexes may help to clarify relationships within the group. Subgrouping within the putative MC group has been relatively shallow. The relatedness of the Trukic languages is apparent even to the non-specialist, as is the relation between Ponapean, Mokilese, and Pingelapese. Among Micronesianists there is a general impression that KSR is most closely related to the Ponapeic languages, but this has never been demonstrated systematically. Perhaps I will eventually manage to do so.

Judith W. Wang
University of Hawaii
December 1981

## NOTES

1. The original version of this paper was written by Kee-dong Lee in May 1972 for a University of Hawaii seminar on historical-comparative Micronesian linguistics conducted by Professor George W. Grace. Since that time Lee has published a *Kusaiean-English dictionary* (Lee 1976) and a *Kusaiean reference grammar* (Lee 1975). However, he never did write the definitive phonology of Kosraean that might have resolved some of the unanswered questions in this paper.

   Judith Wang took up the study of Kosraean in 1977, after Lee was obliged to return to his native Korea, hoping to find the key to the phonology that had eluded Lee. In the interval between Lee's departure and Wang's arrival on the Micronesian linguistic scene, considerable progress had been (and continues to be) made in historical and comparative Micronesian linguistics, some reports of which appear in this volume.

   The present paper is Wang's revision of Lee's original work. Since Wang has not yet succeeded in untangling the synchronic phonology of KSR, this paper must necessarily reflect the analysis presented in Lee 1975. Moreover, owing to limitations of time, additional historical-comparative data not available to Lee in 1972 have not been systematically incorporated, although occasional reference is made to such data.

2. It should be noted, however, that there is some question as to the nature of two of these vowels. Native speakers seem to disagree about the sounds spelled uc and uh. Lee (1975:13-14) indicates that these are distinct vowels, but that uh may alternate with uc under certain conditions. Most words, however, containing one or the other clearly have the one or the other without alternation. The dictionary (Lee 1976) makes no mention of variation between the two sounds.

   However, not all KSR speakers seem to agree with the dictionary. Wang's two principal informants are quite consistent in distinguishing between the sounds spelled uc and uh. For both uc is somewhat higher and has a noticeable diphthongal quality in careful pronunciation. These two speakers generally seem to agree with each other as to which words have uc and which uh. However, their judgments do not wholly agree with those of the KSR speakers who served as consultants for the dictionary. A larger sample of speakers will have to be consulted before any resolution can be expected to this problem.

   The sequence of letters oa actually represents two possibilities in KSR. One clearly is a sequence of two vowels [o:a:] — e.g. toasr [to:a:ṣ] *heavy, sad*. No further attention will be paid to these forms. The other possibility is the troublesome — and interesting — one. For the purposes of the reference grammar Lee treated oa as a vowel among other vowels. However, Lee's own statements (Lee 1975:15-16) indicate that oa appears to be a diphthong rather than a unitary vowel. The behaviour of Wang's KSR consultants appears to be consistent with oa being a diphthong: only one speaker ever pronounced it as a unitary low back rounded vowel, and did so only as a careful (elicitation) pronunciation. Both this speaker and one of the others alternated between a careful diphthong pronunciation [oa] and a casual pronunciation in which the first part of the diphthong appears to have been lost [a:]. A third speaker consistently gave a low back unrounded vowel [a:] and only evidenced awareness that this vowel might be rounded for some other speakers when questioned directly about it. See Section 2.9. for further speculation on this question.

3. For the purposes of this paper Micronesian (MC) means nuclear Micronesian as defined by Bender (1971).

4. Lee (1975) distinguishes between a dental lateral /l/ and /r/, which is described as a palatal flap. Wang finds no evidence that KSR r is ever pronounced as a flap by native speakers. Rather, it is a voiced, retroflexed palatal continuant, more strongly retroflexed than in most varieties of American English, but not unlike the Beijing (Mandarin) r.

5. This n reflex of POC *t occurs in only one form, KSR ninac *mother* (< POC *tina), and is most plausibly explained as a sporadic change.

6. *d and *(dr) are regarded as liquids for the purposes of this paper on the basis of their reflexes in KSR. The KSR reflex of *nd is not a liquid but a fricative sr, with the same strong retroflexion as KSR r.

7. *R is not reflected as a liquid in KSR; however, it is included among the liquids because it appears to be reflected as some variety of r or other liquid in some other OC languages.

8. The alternation in the stem vowel of this and other so-called 'inalienable' nouns (i.e. those for which possession is marked by the direct suffixation of the possessive marker, rather than by a following 'classifier' marked for possession) appears to be morphologically conditioned in the synchronic grammar. It is possible that such alternations were at one time conditioned by the final vowel of the possessive suffix, but this conditioning factor was lost when a (historical) rule of final vowel deletion came into the language. There is no synchronic evidence for the retention of historical vowels, even in underlying forms. The variation one is apt to encounter within a single possessive paradigm, as well as the competition between alternative paradigms for the same noun, both testify to the present fluctuating state of KSR morphophonemics.

9. This may be the same formative as in kihfihnte *saltwater* (< kof *water* + ligature (?) + te *sea, salt water*).

10. See Sections 2.8. and 2.9. for discussion of distinctive velarisation and labialisation in KSR.

11. Grace (1969) uses the convention of writing two or more segments within parentheses to indicate that there is evidence for each of the alternative reconstructions, but that the evidence does not appear to favour one over the other(s).

12. Note also PUL ŕ, cc, STW rh, ch, CRN rh, tch < PMC *t', which also happens to be the source of KSR sr. Also CRL geminate sch → tch, and WOL geminate sh → cc (< PMC *t').

13. Ward Goodenough (personal communication) points out that the present-day (native) Kosraean population is probably mostly descended from a small group of Christian converts who managed to survive the demographic disasters of the 19th century. An 1880 census count showed approximately 200 (adult?) native Kosraeans, as compared to early European contact estimates of several thousand. The people who survived the various European diseases probably did not reflect the original (precontact) dialect distribution of Kosrae. If, as Goodenough suggests, these survivors formed a single Christian convert community, there is every likelihood that the speech of that mixed community would have undergone dialect mixing and levelling. The considerable variation in present-day KSR, as well as the lack of systematic dialect distinctions, may well be the result of such a situation.

14. Two asterisks (**) preceding a form indicate that it is not attested.

15. The order of the rules here is quite arbitrary. There is no evidence as to which order these developments occurred in historically.

16. The same one of Wang's informants who consistently unrounded oe also unrounded the stop in the word intoe *to look*. When shown the written word, this speaker did not recognise it at first and commented that he did not know that there was supposed to be an o there. No trace of lip-rounding was observed when this speaker was asked to pronounce this word, but his pronunciation of other words containing the orthographic sequence of a consonant followed by oac consistently included the phonetic sequence [wɛ] and the preceding consonant appeared to be rounded.

17. Furthermore, a similar problem may exist with respect to the distinctively velarised consonants so glibly summed up in the preceding section (2.8.). That is, it may be that these ought to be analysed synchronically as an underlying plain consonant (which becomes non-distinctively velarised) preceding a high back (central?) unrounded glide. The diachronic version of this analysis would be that the back quality of the vowel was separated from the vocalic nucleus through diphthongisation and that the glide (at least the backness represented by the glide) was subsequently assigned to the preceding consonant. The apparent association between low-vowel dissimilation and distinctive velarisation may be explainable as a transfer of the quality of backness from the vowel to the preceding consonant, with simultaneous or subsequent raising of the now-front vowel. Even the case of twem *(vt) to sew* (< POC *tasimi), while evidently not involving low-vowel dissimilation, still may be analysed as a transfer of the quality of backness from the *a to the preceding consonant.

    All this is pure speculation, of course, that would have to be fitted somehow in a coherent analysis of KSR phonology before it can be used to account for anything.

18. This KSR form probably ought to be attributed to POC *muta(?) plus the *-i transitive suffix, which may have conditioned a different set of changes before it was lost through final vowel deletion.

19. The difference between this and the form woht (also attributed to POC *muta(?)) is that wihte is transitive and properly ought to be attributed to POC *muta(?) + -i.

20. The *(ou) in the reconstruction indicates that the available reflexes of this proto-form do not point unequivocally to either of these vowels, merely that it must have been one or the other of them. KSR seems to point toward *o (see Section 3.1.4.).

21. As pointed out in Note 2, there appears to be some confusion among native speakers regarding the vowels spelled uh and uc. The uh spelling is the one given in the dictionary; the uc spelling was offered by one of Wang's KSR consultants.

22. KSR uh in words like muhta- (< POC *mata *eye, face*) and tuhma- (< POC *tama *father*) does not seem to be so much the result of historical low-vowel dissimilation as of synchronic stress patterns. Both of these words are stressed on the second syllable, suggesting that the uh in the first syllable may be the result of vowel reduction.

23. The word rules is used here solely for reasons of convenience. It is not meant to imply any particular theoretical position with regard to the nature of rules and how (or even if) they should be formalised.

24. See Section 1.3. or Lee (1975:30-32) regarding (synchronic) vowel length in KSR.

25. See Note 22.

26. See Section 2.9.

27. Obviously, these derivations are merest speculation, serving only to sketch out possible paths of development from the reconstructed POC form to the currently-attested KSR form.

28. Presumably final vowel deletion did not apply to monosyllables, or if it did, it produced short vowels that were then lengthened again by monosyllabic lengthening.

29. There is an equally plausible alternative derivation for at least the forms te (< *tansi(k)) and e (< *api), involving *a-raising, loss of *ns/*p, and final vowel deletion. That derivation will not work for acn (< *papine), however, because the *i that would condition *a-raising is not in word-final position and would require an additional rule to delete it. It was to account for forms like *papine and *tasimi (> twem) that the rule of monophthongisation was proposed.

30. If this form is not a borrowing.

31. See Lee (1974) for a description of these and other verbal derivation rules which also play a role in the evolution of the KSR forms under consideration. One would expect that there would be a corresponding set of synchronic phonological rules, but the phonology has resisted all attempts to date at systematic description.

32. This presupposes the existence of a distinct nuclear MC subgroup within Oceanic. However, the group described in Bender (1971) has never been clearly defined in terms of a set of exclusively shared innovations.

## POC RECONSTRUCTIONS

Most of the reconstructions listed below are taken from Grace 1969. Those taken from Blust 1972 are prefixed with a B, those inferred from Dyen 1949 with a D.

| POC | KSR | |
|---|---|---|
| D*anitu | inut | *ghost, spirit* |
| *?anusi | acni | *to spit* |
| B*aŋin (also *nsaŋi) *yellow* | eng | *wind* |
| *aŋoaŋo | ahng | *kind of plant: turmeric* |
| *api | e | *fire* |
| *?aRu(s) | acsr | *current, stream* |

| | | | |
|---|---|---|---|
| *?ate | | esyac-/acsyac- | *liver* |
| *?atoluR | | ahtro- | *egg* |
| *da?a(n) | | lah | *branch* |
| *ndanu(m) | *fresh water, bathe, mix with liquid* | sroano- | *liquid* |
| *ndau(n) | | sra | *leaf* |
| *doŋo/*ndoŋo | | lohng | *to hear* |
| *dua | | lo/luo/lu- | *two* |
| *inu(m) | | nihm | *(vt) to drink* |
| *kali | | kuhlkuhl | *(vi) to dig* |
| *kianto | | kiyacs | *outrigger boom* |
| *ŋkinit | | kihnis | *(vt) to pinch* |
| | | kihnkihn | *(vi) to pinch* |
| *kinta | | kuht | *we (1 pl)* |
| *kuli(t) | | kolo-/kulu- | *skin, bark* |
| *kuRita | | koet | *octopus* |
| *kutu | | kut | *louse* |
| *laŋi(t) | | lucng | *sky, heaven* |
| *laŋo | | loang | *fly (insect)* |
| *lima | | luhm-/lime- | *five* |
| D*madama | | mahlwem | *moon* |
| *manu(k) | | won | *bird* |
| *manipi(s) | | mihnini | *thin* |
| *mapo | | mah | *heal* |
| *mata | | muhta-/mwet | *eye, face* |
| *ŋmata | | wet/wat | *worm* |
| *matolu | | mahtol | *thick* |
| *matudu(R) | | mutul | *to sleep* |
| *ma?udi(p) | | moul | *alive, life* |
| *-mu | | -m | 2 sg poss suff |
| *muta(?) | | wihte | *(vt) to vomit* |
| | | woht | *(vi) to vomit* |
| *napo | *surf* | noa | *wave* |
| *natu | | nahtuh-/nuhtV- | *child* |
| *niu(R) | | nu | *coconut* |
| *ŋapulu(?) | | si-nguhul | *ten* |
| *onom | | on-/ohn- | *six* |
| *pai- | | a- | reciprocal pref |
| *paka- | | ahk- | causative pref |
| *pakiwak | | pahko | *shark* |
| *panua | | acn | *land* |
| *paŋ(ou)(n) | | pahngon | *(vt) to call* |
| | | pahngpahng | *(vi) to call* |
| *papine | | acn | *woman* |
| *patu | | yot | *stone* |
| *mpaya | | pa | *bait* |
| *mpeŋka | | fak | *bat, flying fox* |
| *pi(dr)i | | ir | *twist* |
| *pitu | | it- | *seven* |
| *pitu?u/*pitu?o(n) | | itu/itih | *star* |
| *mpo | | fo | *smell* |
| B*mponot/*ponot | | fonos | *(vt) to block, plug* |
| | | fohnfohn | *(vi) to block, plug* |
| *mpoŋi | | fong | *night* |

| | |
|---|---|
| *mpou | |
| *mpule | |
| *mpulu(t) | *gum, sap, pulp, glue, sticky* |
| *punu(?) | |
| *puti/D*punti | |
| *mputo | |
| *Ruŋma(?) | |
| *sae(t)/*sai/*sei | |
| *nsai | |
| *nsakaRu | *reef* |
| *sake/*nsake | |
| *nsaŋi (also *aŋin) | |
| *si(dr)i(t) | |
| *sili | *to enter* |
| *nsiwa | |
| *sulu | |
| *susu | |
| *taku | |
| *tali | |
| *tama | |
| *tanom/*tanum | |
| *taŋi(s) | |
| *tansi(k) | *sea, saltwater, salt* |
| D*tasimi | *sharpen it* |
| *tia(n) | |
| *tina | |
| *tolu | |
| *topu | |
| *tudu(?) | |
| *tuki | |
| *tu?u(d) | |
| *?ulo(s) | |
| *?uluŋa | *pillow, head rest, to rest one's head* |
| *?uma | |
| *?umu | |
| *?una(p) | |
| *?utup | *to flood, draw water, fill with liquid, soak* |

| | |
|---|---|
| fo | *post, pole, stick* |
| ful | *kind of shell: cowry* |
| fulus | *(vt) to paste, gum* |
| fulful | *(vi) sticky, gummy, adhesive, viscous, juicy, sappy* |
| uni | *to hit, kill* |
| usr | *banana* |
| fuht/fihtac- | *navel* |
| yuwac-/iwac- | *house, shelter, place* |
| lohm | *house* |
| se | *(vt) to tear* |
| suc | *who?* |
| tuhka | *island* |
| -yak/-ack | dir suff: *up, upward* |
| eng | *wind* |
| iri | *(n) masturbation* |
| irii | *(vt) to masturbate* |
| il | *to come, go, move* |
| yuh | *nine* |
| sul | *torch* |
| titi- | *breast* |
| tok/tohkoh-/tuku- | *back (anat)* |
| sucl | *string, rope* |
| tuhma- | *father* |
| taun | *(vt) to bury* |
| tatuh | *(vi) to bury* |
| toan | *(vt) to press, bury* |
| toatoa | *(vi) to press, bury* |
| tuhng | *to cry* |
| te | *beach, seaside* |
| twem | *(vt) to sharpen* |
| twetwe | *(vi) to sharpen* |
| (in-)siyac- | *belly* |
| ninac | *mother* |
| tol- | *three* |
| tuh | *sugarcane* |
| tul | *drip, drop* |
| tuk | *(vt) to pound* |
| tuktuk | *(vi) to pound* |
| tu | *to stand* |
| ulac | *maggot* |
| ilung | *(vt) to prop, support, rest* |
| ilul | *(vi) to rest, prop* *(n) pillow* |
| imac | *garden, field* |
| um | *earth oven* |
| unac-/une- | *fish scale, feather, fur, body hair* |
| uti | *(vt) to fetch water, draw water* |
| ut | *(vi) to fetch water, draw water* |

| | | |
|---|---|---|
| *waka | okah | *root* |
| *walu | oal- | *eight* |
| *waRo | ah | *string, line, rope* |

## BIBLIOGRAPHY


BENDER, Byron W.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 426-465. The Hague: Mouton.

BIGGS, Bruce

1965 Direct and indirect inheritance in Rotuman. In G.B. Milner and Eugénie A. Henderson, eds *Indo-Pacific linguistic studies*, part 1: *Historical linguistics*, 383-415. Amsterdam: North-Holland. (= *Lingua* 14:383-415.)

BLUST, Robert A.

1972 Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list. *Working Papers in Linguistics, University of Hawaii* 4/1:1-43.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:429-436.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics* 1/2:39-84.

HARRISON, Sheldon P.

1978 Transitive marking in Micronesian languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 1067-1127. *Pacific Linguistics*, C-61.

LEE, Kee-dong

1974 Kusaiean verbal derivation rules. Doctoral dissertation, University of Hawaii, Honolulu.

1975 *Kusaiean reference grammar*. Honolulu: University Press of Hawaii.

1976 *Kusaiean-English dictionary*. Honolulu: University Press of Hawaii.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

SOUTH PACIFIC COMMISSION

1981 *Economic development statistical survey*. Noumea.

WANG, Judith W.

1981 Combined Oceanic reconstructions with English finderlist (preliminary version). Computer printout. Department of Linguistics, University of Hawaii.

WILSON, Walter Scott

n.d. Outline of Kusaiean history. MS. Pacific Collection, University of Hawaii library.

# OBJECT MARKING IN MARSHALLESE

Byron W. Bender

This paper focuses on two formatives /ey/ and /iy/ and the multiple functions they fill in contemporary Marshallese (MRS), which include both the formation of transitive verb stems, and the marking of third person objects of such verbs. The variety of ways in which they fill these functions (together with alternative means of transitive verb formation) serve to establish seven major classes of transitive verbs, each of which will be looked at in some detail. Where possible, knowledge of the history of the language will be used to bring understanding to the present complex situation.[1] Although most of the discussion will deal with transitive verbs, it will be extended to include three prepositions which are thought to have had a verbal history, and which demonstrate parallel object marking. The object and absolute pronouns will also be treated because of their close relation to object marking.

The object and absolute pronouns are largely identical, differing in form only in the first and second persons singular. The two sets may be displayed as follows:

| PERSON AND NUMBER | ABSOLUTE PRONOUNS | ABSOLUTE AND OBJECT PRONOUNS | OBJECT PRONOUNS |
|---|---|---|---|
| 1s | n̄a /gah/[2] | | eō /yęh/ |
| 2s | kwe /qey/ | | eok /yęq/ |
| 3s | | e /yęy/ | |
| 1pi | | kōj /kęj/ | |
| 1pe | | kōm /kęm/ | |
| 2p | | koɱ /qęɱ/ | |
| 3p | | er /yęr/ | |

The plural pronouns as listed may be used for any number of referents greater than one, but may also be further specified through the addition of suffixes as dual (-ro /-rew/), trial (-jel /-jęl/ ∿ -jeel /-jęyęl/), quadruple (-men /-men/ ∿ -eaṉ /-yag/), or multiple (-wōj /-węj/). Both sets serve only for human referents; there are no full pronouns in the language for non-human referents. The latter have as their only object substitute forms the two formatives under discussion.

The absolute pronouns may be used in equational sentences such as the following, where the object pronouns may not be substituted.

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 443-465. *Pacific Linguistics*, C-80, 1984.

1a. Rukaki n̄a. 1b. *Rūkaki eō.
*teacher* 1sABS 1sOBJ

*I'm a teacher.*

On the other hand, the absolute pronouns may generally be substituted for the object pronouns in the positions they typically occupy, following transitive verbs as their objects, or as objects of prepositions such as n̄an /gan/ *to*, jān /jan/ *from*, or kōn /kẹn/ *with, concerning, because of*.

2a. Kwō-n jab deṉ̄ḷọke eō.
2sSM-IMP *not hit* 1sOBJ

*Don't hit me.*

2b. Kwō-n jab joon n̄a.
2sSM-IMP *not press down on* 1sABS
*Don't put your weight on me.*

3a. Kwō-n kaiur bwe e-naaj boṉ̄-e eok.
*hurry because* 3sSM-FUT *night*-TM 2sOBJ

*Hurry up before it gets dark.*

3b. E-maroṉ̄ Jeptōm̧ba-ik kwe ŋa ān-in.
3sSM-*able September*-TM 2sABS LOC *islet*-DEM: *this*

*September might still see you on the island.*

4a. E-ar jab eḷtok n̄an eō.
3sSM-PAST *not pay attention to* 1sOBJ

4b. E-ar jab eḷtok n̄an n̄a.
1sABS

*He didn't pay attention to me.*

While there is a decided preference for the object pronouns following transitive verbs in sentences like those of 2 and 3, the absolute pronouns are almost equally common following prepositions as in 4, and any differences in meaning between sentences such as 4a. and 4b. are too subtle to have been clearly identified as yet.

Noun phrases with a variety of internal structures may occur as objects of transitive verbs. Common structures include, in addition to one of the pronouns, a proper noun, or a common noun followed by a demonstrative (with optional additional modifying material). All such noun phrases are definite, and all normally follow a transitive verb directly, although certain adverbial material may intervene. Thus in the following sentence, the object of the transitive verb bwini is the noun phrase ek kā, consisting of the noun ek and the demonstrative kā.[3]

5a. Kwō-n bwin-i m̧ōk ek kā.
2sSM-IMP *count*-OM *please fish* DEM: *these*

*Count these fish please.*

The verb and its object are here separated by an optional adverbial particle, common as a softener in such imperatives. The verb consists of a root /bin/ and an object marker /iy/, one of the forms that are the subject of this paper.

## CLASS I. THEMATIC TRANSITIVE VERBS

The formatives /ey/ and /iy/, together with several others, perform five distinct functions for the various classes of transitive verbs and prepositions. The functions will be introduced with respect to the class of verbs for which they are maximally differentiated in form, Class I Thematic, one of four strong classes. The other classes, strong and weak, will then be introduced and surveyed with respect to the five functions. The terms strong and weak are applied conventionally, with weak classes being those whose membership is still open to new additions according to regular derivation patterns, and strong classes those with closed, fossilised membership. Other terms used to name the classes, such as thematic, will be explained in what follows.

### F-1. Transitive marker and general object marker

The formatives in this function serve to mark a transitive relation between a verb (or a preposition) and its object. Neither /ey/ nor /iy/ fills this function for the thematic classes (which, as will be seen, include not only Class I Thematic, but also Class II Thematic High, and Class VII Weak Thematic). For these classes the formative filling F-1 has the shape VC, of which the vowel will be called the stem vowel, and the consonant the thematic consonant. Both the vowel and consonant of thematic formatives (for the strong classes I and II) were part of the verb root historically, and the transitive stem was formed by a vowel suffixed to the root. When the deletion of final consonants became a general rule in the language, the thematic consonants were lost from unsuffixed forms of the root where they were final, but were retained in the transitive forms where they were 'protected' by the transitivising suffix. Thus, for example, a representative Class I verb /deget/ *hit* was transitivised by the addition of a suffix thought to have been *-a before a 3s noun phrase object: deget-a *hit s..*[4] With final consonant deletion, other unsuffixed forms of the verb came to be dege, while the transitive form remained degeta. Later a rule of final vowel deletion changed both these forms, to deg and deget respectively.

| | UNSUFFIXED FORM | TRANSITIVE FORM |
|---|---|---|
| EARLIER FORM | deget | degeta |
| AFTER FCD | dege | degeta |
| AFTER FVD | deg | deget |

Thus the former transitive marker, a vowel suffix, was lost, and a former portion of the root, the last vowel and consonant, came to be interpreted as the transitive marker. Since the various verb roots that underwent parallel development had a variety of final consonants (and vowels), a number of different thematic consonants (and stem vowels) today serve as transitive markers for these verbs: while -et is the marker for deget, -en serves the same purpose for pewen, and -ak serves for ļemak, to take just two other examples.

| | |
|---|---|
| /pewen/ *lower the sail of* | /pew/ *lower sail; to land* |
| /ļemak/[5] *determine s.* | /ļam/ *scheme, shape* |

(The forms on the right are infinitive counterparts of the transitive verbs on the left.) Each such VC transitive marker must be learned separately for each of these verbs, making them truly 'strong'. There has been some interchange, so that certain verbs today have thematic consonants other than the original

ones, and there has been great generalisation of one marker, -ik, to form a new class, Class VII Weak Transitive, as will be seen.

In addition, the infinitive counterparts of many transitive verbs, including /deget/, have undergone reduplication, thereby increasing the formal difference.

| | | | |
|---|---|---|---|
| /deget/ | *hit s.* | /degdeg/ | *hit, slap* |
| /dibej/[6] | *spear s.* | /dębdęb/ | *to spear* |

Furthermore, a number of Class I verbs manifest no thematic material when compared with their roots or infinitive counterparts. The latter, like the preceding two examples, undergo reduplication, and this today is the sole formal difference. Some may have had thematic material historically which has since been lost or reanalysed; others probably had none. They are subsumed under Class I as a subclass with no extant thematic material because their behaviour with respect to the four functions yet to be introduced generally parallels that of the other members.

| | | | |
|---|---|---|---|
| /kad/ | *throw s.* | /kadkad/ | *throw* |
| /ɱan/ | *kill s.* | /ɱanɱan/ | *kill* |
| /jęm/ | *sharpen s.* | /jęmjęm/ | *sharpen* |

For some, the reduplication in the infinitive is not full, but consists of initial consonant doubling.[7]

| | | | |
|---|---|---|---|
| /dewer/ | *put s. down* | /ddewer/ | *deposit* |
| /jeyeẅ/ | *erase s.* | /jjeweẅ/ | *obliterate* |

Finally, for some there is full suppletion.

| | | | |
|---|---|---|---|
| /kag/ | *eat s.* | /ɱegay/ | *food; eat* |

In general, as will be seen from the discussion of the other verb classes, the criterion by which they are established is the nature of the formal difference between transitive verbs in F-1 and their infinitive counterparts. Subsumed under Class I are a number of verbs for which a thematic VC on the transitive form constitutes the sole difference, others which show reduplication in the infinitives as well, and others that have the latter but not the former, which might be thought of as 'zero thematics'. This varied class has an identity of behaviour with respect to the object marking functions to be discussed which contrasts with that of the other six classes. F-1, the behaviour of stems marked as transitive, will now be exemplified using the Class I verb deñōt /deget/ *hit s.*.

Class I stems marked as transitive fill in F-1 a most general object marking function for any definite noun phrase that overtly follows the verb in the same clause, with or without intervening adverbial material. Such noun phrases may be singular or plural, human or non-human, nouns with accompanying demonstratives, or pronouns. There are no restrictions other than that the noun phrase object be definite and expressed in the same clause following the verb.

6a. Rūkaki eo e-ar deñōt ļadik eo.
*teacher* DEM 3sSM-PAST *boy* DEM: *the (sg)*

*The teacher slapped the boy.*

6b. Rūkaki eo ear deñōt e.
PRO: 3sOBJ *(human)*

*The teacher slapped him.*

7a. Rūkaki eo ear deṇōt ļadik ro.
DEM: *the (pl, human)*

*The teacher slapped the boys.*

7b. Rūkaki eo ear deṇōt er.
PRO: 3pOBJ *(human)*

*The teacher slapped them (human).*

8a. Kōrā eo ear deṇōt kweet eo.
*woman* *octopus* DEM: *the (sg)*

*The woman pounded the octopus.*

9a. Kōrā eo ear deṇōt kweet ko.
DEM: *the (pl, non-human)*

*The woman pounded the octopuses.*

The b. sentences of 6. and 7. have object noun phrases filled by object pronouns that are equivalent to the object noun phrases of the respective a. sentences. The absence of non-human full pronouns does not permit parallel b. sentences for 8. and 9.

## F-2. Transitive stem extension and general object marker

The formative /ey/ may optionally serve to extend Class I stems that have been formed as discussed under F-1. Thus each of the above a. and b. sentences 6. through 9. could equally well have a verb of the form deṇōte /degetey/. Some of the pressures working both for and against this optional function will be noted in later discussion.

## F-3. 3s anaphoric object marker

The formative /ey/ obligatorily marks third person singular objects, both human and non-human, of Class I verbs in anaphoric contexts. This may occur when the referent is understood from the context and unexpressed, or when the noun phrase expressing it has been relativised, or moved to the front of the clause by other processes such as topicalisation.

6c. Ewi ļadik eo rūkaki eo e-ar deṇōt-e?
*where? (sg)* *boy* DEM *teacher* DEM 3sSM-PAST *slap*-OM

*Where is the boy the teacher slapped?*

8c. Ewi kweet eo kōrā eo ear deṇōte?
*octopus* *woman*

*Where is the octopus the woman pounded?*

## F-4. 3p non-human anaphoric object marker

The formative /iy/ obligatorily marks third person plural non-human objects of Class I verbs in anaphoric contexts.

9c. Erki kweet ko kōrā eo ear denōt-i?
*where?* (pnh) *octopus* DEM: 3pnh *woman* *slap*-OM

*Where are the octopuses the woman pounded?*

Parallel anaphoric treatment of third person plural human objects is not possible.[8] The equivalent for sentence 7. would obligatorily use the full pronoun.

7c'. Erri ḷadik ro rūkaki eo ear denōt(e) er?
*where?* (ph) *boy* DEM: ph *teacher* (F-2) PRO: 3pOBJ

*Where are the boys the teacher slapped?*

The following would be an acceptable but rarely used alternative to 6c, where the object referent is to be emphasised.

6c'. Ewi ḷadik eo rūkaki eo ear denōt(e) e?
*where (sg)* DEM: s (F-2) PRO: 3sOBJ

*Where is the boy the teacher slapped?*

The optional F-2 stem extension is possible in both 6c'. and 7c'., as is true for any F-1 context.

## F-5. 3p non-human object marker

The formative /iy/ may optionally be used to mark third person plural non-human expressed objects. Thus F-5 can be seen as the optional extension of F-4 to non-anaphoric contexts, where the noun phrase object is expressed. As such, it is in conflict with F-2 for Class I verbs (and as will be seen, for the verbs of several other classes as well). Parallel extension of F-3 for singular objects presents no such conflict with F-2 and is indistinguishable from it, since the marker in both cases is /ey/.

9d. Kōrā eo ear denōt-i kweet ko.
*woman* *pound*-F-5 *octopus*

*The woman pounded the octopuses.*

Thus Class I verbs may have three alternative forms before third person plural non-human objects: with neither formative (F-1): denōt (as in 9a.); with /ey/ (F-2): denōte; and with /iy/ (F-5): denōti (as in 9d.). F-1 is the preferred alternative, and presents a way out of the conflict between F-2 and F-5, calling as they do for different formatives in the same suffix position. Nevertheless, there are competing pressures favouring on the one hand F-2, and on the other F-5, and for some classes F-1 does not present a way out. More will be said on this later, after all the classes have been surveyed. The formatives filling each function for each class are summarised in Table 1.

## CLASS II. THEMATIC HIGH

One key difference distinguishes Class I and Class II verbs, and that is revealed by the complementarity of their stem vowels: Class II verbs have only the high stem vowel /i/, whereas Class I verbs have only the non-high vowels /e/, /ę/, or /a/. This difference has wide-ranging effects because of a limited vowel harmony that is observed with respect to the /ey/ and /iy/ formatives.

Although /ey/ fills both F-2 and F-3 for Class I verbs, it is excluded from both functions for Class II verbs, which permit only /iy/ to follow their /i/ stem vowel (with intervening thematic consonant). The result then is that /iy/ fills all four functions F-2 through F-5, thereby neutralising the contrast between the marking of singular and plural objects. Otherwise the two classes are quite parallel. F-1 for Class II is generally marked by thematic material, and the relations between transitive verbs and infinitive counterparts parallel those of Class I.

Simple thematic

| | | | |
|---|---|---|---|
| /jagit/ | *cry for s.* | /jag/ | *cry* |
| /jigit/ | *fart on s.* | /jig/ | *fart* |
| /harin/ | *imitate s.* | /har/ | *image* |
| /kibij/ | *dig s.* | /kębˌ/ | *dig* |

Thematic vs. reduplication

| | | | |
|---|---|---|---|
| /m̧ijit/ | *cut s.* | /m̧ijm̧ij/ | *cut* |
| /dapij/ | *hold s.* | /dapdęp/ | *hold* |
| /yigir/ | *sprain s.* | /yigyig/ | *change for worse* |
| /yigit/ | *twirl s.* | /yigyig/ | *kinky* |

Thematic vs. initial consonant doubling

| | | | |
|---|---|---|---|
| /likit/ | *deposited s.* | /llik/ | *contribute* |
| /tawik/ | *mend a net* | /ttaw/ | *mend nets* |

'Zero thematic' vs. reduplication or consonant doubling

| | | | |
|---|---|---|---|
| /lim/ | *fold s.* | /lęmlęm/ | *fold* |
| /liw/ | *scold s.* | /lliw/ | *angry* |

Miscellaneous

| | | | |
|---|---|---|---|
| /kalbin/ | *plant s.* | /kallib/ | *to plant* |
| /kajitikin/ | *question s.* | /kajjitęk/ | *question* |

Suppletive

| | | | |
|---|---|---|---|
| /yilim ∿ yinim/ | *drink s.* | /yidahak/ | *to drink* |

The syntactic behaviour of such verbs parallels that of Class I verbs, except that /iy/ everywhere takes the place of /ey/ in F-2 and F-3.

## CLASS III. STRONG HIGH

These verbs, like those of Class II, have high stem vowels, but differ in that they do not observe the vowel harmony found there. Unlike any other class, they permit the /ey/ formative to follow a high stem vowel. In fact, all of them end obligatorily in this formative, which is thus said to fill F-1. As will be seen, there are a number of indications that this formative was originally part of the verb root for original members of this class. Proto-forms have been reconstructed for several: the forms cited are Proto-Oceanic.[9]

| | | | | |
|---|---|---|---|---|
| /r̈ipey/ | *smash s.* | /r̈ipr̈ip/ | *break* | < *ropa |
| /jinjiney/ | *swear at s.* | /jęnjęn/ | *curse* | < *tina |
| /witey/ | *rain on s.* | /węt/ | *rain* | < *ʔusa |
| /tiley/ | *ignite s.* | /ttil/ | *burn* | < *sila |
| /kiley/ | *recognise s.* | /kalęlkęl/ | *try to recognise* | < *kila |

It will be noted that these proto-forms all end in *a, and with the exception of the first example, are preceded by a high vowel. The final vowel is thought to have been raised to the mid vowel /e/, and — supplied with an epenthetic /y/ — reanalysed as the /ey/ formative. This would seem to be the basis for the formation of this class: a group of roots containing a sequence of high and mid vowels, the latter of which came to be interpreted as the /ey/ formative. Such historic final vowels have been lost in the infinitive counterparts, but participate in a regular alternation that results in a lowering of preceding high vowels to a derivative high-mid quality. This alternation, common in both noun and verb morphology, can be summarised as follows: CiCe- / CęC#. This expected quality is found in three of the infinitives above, as well as in a good majority of the 40 some members of this class that have been discovered to date, of which a few more examples follow.

| | | | |
|---|---|---|---|
| /biney/ | *count s.* | /bęnbęn/ | *count; math* |
| /jibey/ | *hold s.* | /jębjęb/ | *hold* |
| /kiney/ | *compose s.* | /kkęn/ | *invent, compose* |
| /tiwey/ | *dig s. out* | /ttęw/ | *dig taro* |
| /biwey/ | *cut s.* | /biqabęq/ | *cut* |

The last infinitive, with original final low vowel preserved medially in the reduplicative process, exemplifies nicely the supposed history of this class.[10] This is not to say that all present members of the class actually participated in that history. Some no doubt had their final vowels lowered by analogy with true members. This seems to have been the case for the following three verbs, for example, the first of which retains a variant that belongs to Class IV.

| | | | |
|---|---|---|---|
| /jiqjiqey/ (∿ /jiqjiqiy/) | *massage s.* | /jiqjiq/ | *pound breadfruit or taro; massage* < *tuku |
| /kirkirey/ | *tickle s.* | /kirkir/ | *tickle* < *kidi |
| /kenajnijey/ | *curse s.* | /kenajnęj/ | *curse* < *natu |

The last pair of examples introduce a second alternation involving the high-mid vowel, in which the original vowels were in the opposite order from that of the alternation above: CeCi- / CęC#.[11] Derivative high-mid vowels, as in the last infinitive, could thus be interpreted as having two possible sources — historic CiCe or CeCi — and are often reinterpreted counter to actual history with the today more popular CiCe. This can be shown to have occurred for a number of words outside this class as well. The verb /kenajnijey/ has an alternative /kenajnęjęy/ that reflects neither alternation pattern, but is formed instead by treating the infinitive as basic and adding /ey/, which is then raised to /ęy/ by a rule discussed under Class V termed the 'high-mid harmony' rule. Such alternates should be considered members of Class V; there are at least a dozen parallel examples, a few of which are given here.

| Class III | | Class V | | Infinitive | |
|---|---|---|---|---|---|
| /yamijey/ | ∿ | /yamęjęy/ | *mourn s.* | /yamęj/ | *mourn* |
| /bawrikey/ | ∿ | /bawrękęy/ | *to plug s.* | /bawręk/ | *plugged* |
| /hawelikey/ | ∿ | /hawęlękęy/ | *sting s.* | /hawęlęk/ | *Portuguese man-o'-war* |
| /yaklipey/ | ∿ | /yaklępęy/ | *rob a hen of eggs* | /yaklęp/ | *rob eggs* |
| /yakpiley/ | ∿ | /yakpęlęy/ | *jettison s.* | /yakpęl/ | *lighten ship* |

Another possible source for members of Class III was verbs in which an earlier non-high vowel was somehow raised, as seems to be true for /r̈ipey/ from *ropa. This raising is not well understood, but can be shown to have occurred elsewhere as in /mijak/ *fear* from *mataku, or /yamijey/ *mourn* from *pa-mate.

As might be expected in view of the high vowel harmony prevailing elsewhere, there are strong pressures to bleed members from Class III, and a number of variants confirming this (in addition to those like /kenajnęjęy/ are to be found, including a variant of one of the verbs cited above: /biniy/ *count s.*. For this verb the infinitive /bęnbęn/ still serves as a reminder of history; other infinitives have given way, as for example /ttil/ *burn*, or the variant for /jęnjęn/ *curse*: /jinjin/.

Since /ey/ now fills F-1 as part of the basic transitive stem for this class, it is precluded from distinctively marking F-2 and F-3. It is replaced by /iy/ in F-4 and F-5 for plural objects, one of the reasons for analysing it today as a transitivising increment rather than an integral part of the verb root — as is the case in the next class to be considered.

## CLASS IV. STRONG ZERO

This is an extremely small class, with a total of six members discovered to date. They divide into two subclasses of three each. The members of Class IVa are monosyllabic in F-1, F-2, and F-3, and this one syllable contains an /ey/ formative. Whether or not to identify this as the /ey/ under discussion is problematic. Favouring the identity is the fact that a second /ey/ is not added for F-2 and F-3. Countering it is the fact that /iy/ is added to /ey/ — instead of replacing it — in F-4 and F-5. The latter view is adopted here for its apparent historic validity.

| F-1, F-2, F-3 | F-4, F-5 | | Infinitive | |
|---|---|---|---|---|
| /jey-Ø/ | /jęyiy/ | *write s.* | /jęyjęy/ | *to write* |
| /mey-Ø/ | /męyiy/ | *chew s.* | /męymęy/ | *to chew* |
| /qey-Ø/ | /qęyiy/ | *scratch s.* | /qęyqęy/ | *to scratch* |

The root vowel for these verbs can be said to be the mid /e/. The raising in the F-4, F-5 forms takes place generally when /e/ precedes /i/ with any of the semiconsonants /y/, /w/, or /h/ intervening. The fact that this derivative quality is also found in the infinitives points toward an earlier high vowel there, e.g. *jeyijeyi. The F-1, F-2, F-3 forms may be derived from an earlier form that included the transitivising suffix *-a (already referred to in connection with the development of thematic consonants): *jeyi-a. The i's have been dropped from forms like the latter, yielding *jeya, and the final a's were later lost by final vowel deletion. Harrison (1977:50-51) has demonstrated how the dropping of such i's might have come about through the reinterpretation of the paradigms of some roots with final high vowels as being consonant-final instead. Parallel developments have taken place in Kiribati (KIR) and Mokilese (Harrison 1977:8), and in Ponapean (Ken Rehg: personal communication).

Class IVb consists of three roots that are compounded with one of the three directional postpositions: tok /teq/ *hither*, ļǫk /ļaq/ *thither*, and waj *toward the speaker*. The first two listed are compounded obligatorily, the third optionally.

| F-1 | F-2, F-3 | F-4, F-5 | |
|---|---|---|---|
| /ley-DIR-Ø/ | /ley-DIR-(ey)/ | /liy-DIR-(iy)/ | *give* |
| /jew-DIR-Ø/ | /jew-DIR-(ey)/ | /jiw-DIR-(iy)/ | *throw* |
| /bek-(DIR)-Ø)/ | /bek-(DIR)-(ey)/ | /bik-(DIR)-(iy)/ | *bring, take* |

The infinitive forms are identical to those for F-1. The fact that the directionals are only optionally compounded with /bek/ makes it possible for them to occur instead in their usual post position, following the /ey/ or /iy/ increments of the uncompounded stem. Thus, for example, /bek/ in F-4 with the directional tok may be alternatively:

10. Erki peba ko kwa-ar
*where? paper pl* 2sSM-PAST

{ bū̄ktok /bik-DIR-Ø/ ; bū̄ktoki /bik-DIR-iy/ ; bū̄ki tok /bik-iy # DIR/ } ?

*Where are the papers you brought?*

Two interrelated points are especially noteworthy concerning the forms of these verbs. First, each has a mid stem vowel in F-1, F-2, and F-3 which is replaced by a high vowel in F-4 and F-5, and secondly, this replacement alone is sufficient to signal the latter functions, just as the presence of the mid vowel signals the former. This is the only class (Class IVa and IVb) in which either the /ey/ or /iy/ formative is not mandatory in F-3, and Subclass IVb is the only one in which /iy/ is not mandatory in F-4.

## CLASS V. WEAK

This and the following two classes represent productive means of transitive verb formation. The infinitive counterparts of Class V verbs end in full consonants (as opposed to semiconsonants) that are preceded by vowels other than high. The transitive stem is formed by the addition of the /ey/ formative in F-1, and the distinctive use of this same formative is thereby precluded for either F-2 or F-3. It is replaced by /iy/ in F-4 and F-5.

| | | | |
|---|---|---|---|
| /kateqey/ | *sacrifice s.* | /kateq/ | *sacrifice* |
| /ɱeṇ̃ey/ | *trick s.* | /ɱeṇ̃/ | *lie, deceive* |
| /beqey/ | *throw sand on s.* | /beq/ | *sand* |
| /nęnęy/ | *put a mulberry poultice on s.* | /nęn/ | *mulberry* |

Following are some loans from English.

| | | | |
|---|---|---|---|
| /bahyibe⌊ey/ | *teach s. Bible* | /bahyibe⌊/ | *Bible* |
| /jewebey/ | *put soap on s.* | /jeweb/ | *soap* |
| /tęyępęy/ | *tape s.* | /tęyęp/ | *tape* |
| /męyę⌊ęy/ | *post s.* | /męyę⌊/ | *mail* |

The last two examples, and the last one in the first group above, illustrate the 'high-mid harmony' rule which prevents the mixing of mid and high-mid vowels within a word and in this case raises the vowel of the /ey/ suffix to agree with the stem vowel(s). Class V, like the other two weak classes to follow, is a large one for which examples could be extended indefinitely.

## CLASS VI. WEAK HIGH

Like Class V, the infinitive counterparts of verbs in this class end in full consonants, but in this case the vowel preceding the consonant is always high. Thus there is the same sort of complementarity between Classes V and VI that was noted between Classes I and II — except that here it is root vowels rather than stem vowels, with the harmonising formative ( /iy/for Class IV)

being added to form the stem. The net effect is complete neutralisation among all the functions: for Class VI, and for Class VI alone, one and the same transitive stem ending in /iy/ fills all five functions.

| | | | |
|---|---|---|---|
| /kilbï̤iy/ | *cover s. up* | /kilbï̤/ | *cuddle; blanket* |
| /m̧akï̤iy/ | *prick s.* | /m̧akï̤/ | *pandanus thorns* |
| /kim̧ï̤iy/ | *rub herbs on s.* | /kim̧ï̤/ | *herb ointment* |
| /jimagigiy/ | *ferment s.* | /jimagig/ | *fermented toddy* |

Following are some loans from English.

| | | | |
|---|---|---|---|
| /jiwibiy/ | *make s. into soup* | /jiwib/ | *soup* |
| /jipijiy/ | *lecture s.* | /jipij/ | *speech, lecture* |
| /tanijiy/ | *dance with s.* | /tanij/ | *dance* |
| /wiyiniy/ | *win a contest* | /wiyin/ | *win* |
| /liwijiy/ | *lose a contest* | /liwij/ | *lose* |
| /bawiniy/ | *weigh s.* | /bawin/ | *pound, scales, weigh* |

## VII. WEAK THEMATIC

These verbs are derived from roots that end in one of the semiconsonants: /y/, /w/, or /h/, which may be preceded by any of the vowels. To such roots is added in all cases for F-1 the transitive marker /(y)ik/. The /ik/ of this marker can be seen as paralleling in canonical form the thematic markers that fill F-1 for Classes I and II; hence the word thematic in the name of this class. The presence of the /y/ and its optionality suggest an earlier particle status for the marker, during which time the /y/ was added prothetically — a fairly common means of fulfilling the canonical requirement that all free forms begin and end with a (semi)consonant. It is even possible that /yik/ may have metathesised from an earlier /kiy/ (<*ki), as can be shown to have happened for other particles such as the construct /yin/ (<*ni) and the complementiser /bey/ ∿ /yeb/ (<*be <*ba). The high stem vowel /i/ of the marker triggers the same limited vowel harmony observed in Classes II and VI, so that /iy/ alone fills F-2 through F-5 in the same non-distinctive way it does for Class II, the strong Thematic High class.

Many native members of the class rarely or never occur with /y/ in the marker.

| | | | |
|---|---|---|---|
| /biwik/ | *shoot s.* | /biw/ | *gun* |
| /kahihik/ | *make faces at s.* | /kahih/ | *pout* |
| /m̧iyik/ | *dirty one's diapers* | /m̧iy/ | *soft faeces; defecate* |
| /yawik/ | *tattoo s.* | /yaw/ | *tattoo* |

Others, especially derivative causatives and recent loans, show it optionally.

| | | | |
|---|---|---|---|
| /karittew(y)ik/ | *deflower s.* | /rittew/ | *old* |
| /kawinew(y)ik/ | *dye s.* | /winew/ | *paint; colour* |
| /kiļiw(y)ik/ | *glue s.* | /kiļiw/ | *glue* |
| /jikiriw(y)ik/ | *fasten with screws* | /jikiriw/ | *screw, bolt* |

Generally, its presence or absence seems to be a function of lexicalisation (being included in new derivatives) and of tempo (dropped in more rapid speech).

This class could be termed 'pseudo-thematic' because of the way in which it parallels the strong thematic classes, specifically Class II Thematic High. In

fact, one possible account of the origin of this class would have some members being former members of Class II: those like /m̩iyik/ above, which happened to have a semiconsonant as the penultimate consonant, /i/ as stem vowel, and /k/ as final consonant. After final consonants became thematic, with the reanalysis implied, some interchange of thematic consonants took place, with /k/ being favoured as replacement for others. A nucleus of candidates for new class membership was thus formed, together with a marker /(y)ik/ that could be used by others of different canonical (and class) origins: specifically, roots that happened to have a semiconsonant preceding the final vowel: ...CVYV# (where Y stands for semiconsonant). When final vowel deletion left such roots with final semiconsonant, their resemblance to infinitives of nuclear members (e.g. /m̩iy/) was sufficient for the /(y)ik/ marker to be extended for their inclusion.

Another possible account would have vowel-final roots of the latter sort transitivised by a separate particle /kiy/, which with final vowel deletion affecting such roots, metathesised to /yik/ and proceeded toward a closer union with them, as suffix. These two possible accounts are not mutually exclusive, and Class VII may have been born as a result of the merger of both developments. The second may have influenced the first: the /k/ of the particle may have contributed to replacement of historic thematic consonants by /k/, or conversely, the incipient /(y)ik/ marker may have encouraged the metathesis and union of the particle /kiy/. This latter could have taken place under the same influences posited by Harrison (1977:39-40) for the apparent metathesis of the 1s and 2s object pronouns: /yẹh/ < /hẹy/ < *ai and /yẹq/ < /qẹy/ < *ko, respectively.

## PREPOSITIONS

The formatives /ey/ and /iy/ perform some functions for prepositions that parallel those performed for verbs. Cognates of the first three prepositions listed below in some other Micronesian languages are still partially verbal in function, and are thought to have originated as verbs. The last listed has not been attributed a verbal history, but the present functions of /ey/ with it may have been extended by analogy. They are listed with the major case roles they mark for their noun phrase objects.

| | | |
|---|---|---|
| jān | /jan/ | SOURCE |
| n̄an | /gan/ | GOAL, BENEFACTIVE |
| kōn | /kẹn/ | INSTRUMENT, COMITATIVE, CAUSE |
| i | /yiy/ | LOCATIVE |

The filling of the functions is identical for the first two (see Table 1); the last two each present separate departures from what might be seen as the overall preposition pattern. Part of this pattern is for F-1, F-2, and F-5 to be filled by a zero marker. One might be tempted to say that F-2 and F-5 are not filled for prepositions — rather than to say that they are filled by zero — but the fact that they are filled for /kẹn/ indicates that they are potentially fillable for all. Sentences 4a. and 4b. illustrate F-1 with /gan/. For it and /jan/, F-3 is filled by /ey/ and F-4 by /iy/. /kẹn/ is deviant in having the partially suppletive /kahkey/ and /kahkiy/ in these functions, and by extending the same forms respectively to F-2 and F-5. /yiy/ deviates from the prepositional pattern by having a single form, with a compromise /ẹy/ — neither /ey/ nor /iy/ — in both F-3 and F-4. This formative, however, would have to be interpreted as basically /ey/, with its vowel raised to high-mid by the preceding /yiy/ stem. The most striking feature of the deviancy is that

Table 1: Formatives filling transitive and object marking functions in present-day Marshallese

| FUNCTIONS | TRANSITIVE VERB CLASSES WITH REPRESENTATIVE EXAMPLES AND INFINITIVES | | | | | | | | PREPOSITIONS | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | II | III | IVa | IVb | V | VI | VII | SOURCE/ GOAL BEN | INSTR COM CAUSE | LOC |
| | Thematic | High Thematic | Strong High | Strong Zero | | Weak | Weak High | Weak Thematic | | | |
| | deget degdeg | ṃijit ṃijṃij | biquy biqabẹq | jey jẹyjẹy | ley-DIR ley-DIR | beqey beq | tiř̈iy tiř̈ | biwik biw | jan/ gan | kẹn | yiy |
| F-1 (OBL) Transitive Marker | {e/a}C | iC | -ey | [ey] Ø | [e] Ø | -ey | -iy | -(y)ik | Ø | Ø | Ø |
| F-2 (OPT) Stem Extension | (-ey) | (-iy) | [ey] | [ey] Ø | [e](-ey) | [ey] | [iy] | (-iy) | Ø | (→kahkey) | Ø |
| F-3 (OBL) 3s Anaphoric Object Marker | -ey | -iy | [ey] | [ey] Ø | [e](-ey) | [ey] | [iy] | -iy | -ey | →kahkey | -ẹy |
| F-4 (OBL) 3p Non-human Anaphoric OM | -iy | -iy | →iy | [ḗy]-iy | [→i](-iy) | →iy | [iy] | -iy | -iy | →kahkiy | -ẹy |
| F-5 (OPT) 3p Non-human Object Marker | (-iy) | (-iy) | (→iy) | [ḗy](-iy) | [→i](-iy) | (→iy) | [iy] | (-iy) | Ø | (→kahkiy) | Ø |

KEY: ( ) enclose optional formatives

[ ] enclose formatives identical to stem (or root) formative, thus indicating that formative does not perform function distinctively

→ replacive or suppletive for F-1 stem material

----- indicates /iy/ 'isogloss'

this is the only place where /ey/ — the usual filler of F-1, F-2, and F-3 except when supplanted by /iy/ as a result of vowel harmony — infringes on the domain of /iy/ as filler of F-4.

## THE PRESENT SYSTEM

The complex pattern of the filling of the functions for all transitive verb classes and the prepositions is summarised in Table 1. For Classes I, III, and V, and for two of the prepositions, /ey/ serves to mark singular objects in F-3, in opposition to the marking of plural objects in F-4 (and F-5) by /iy/. This clear singular-plural correlation is broken by the intrusion of /iy/ into /ey/ territory for Classes II, VI, and VII, as shown by the broken line in Table 1. A more ideal marking system would have the line run straight across the table between F-3 and F-4. There seems little likelihood, however, that there will be any move in this direction, because two of the three classes where the intrusion takes place are weak productive classes, and the principle of high-vowel harmony is strictly observed in assigning new full-consonant-final roots to Class VI vs. V. The principle is also at work bleeding verbs from Class III, and seems firmly entrenched on all sides. Because it at the same time provides a clear domain for the /ey/ marker in the addition of new verbs to Class V, there is little likelihood that /ey/ will be completely replaced by /iy/.

As will be seen, some classes (such as III and V) received the /ey/ marker as part of their historic endowment, and thus came ready-made with the stem extensions of F-2. Class VI received the /iy/ marker in the same way. Other classes (such as I, II, and VII) had neither historically, inheriting instead what might be viewed as a zero following their final consonants in F-1. These latter classes have since taken on one of these two markers (as appropriate in keeping with high-vowel harmony) as stem extensions by analogy with the former. Some classes thus have obligatory stem extensions, while others do not, but tend to acquire them in the general transitive/object marking function, oblivious to singular-plural distinctions. This can be seen as the motivating force behind F-2: that all transitive stems, in their general non-anaphoric object marking capacity, should end in either /ey/ or /iy/.

Stems in their F-1 forms may be freely used before all types of expressed objects, including non-human plural. Extended stems in F-2 should be equally free. The /ey/ used to extend stems of Classes I and IVb (and the preposition /kẹn/ as /kahkey/), however, clashes with the /iy/ called for by F-5, which is sensitive to the plurality of expressed non-human objects. The obligatory /ey/ in F-1 and F-2 for verbs in Classes III and IV similarly clashes with the replacive /iy/ called for in their F-5 forms. These opposing tendencies, the one oblivious to non-human object number, the other sensitive to it, result in a state of variation between /ey/ and /iy/ for verbs in all classes that do not observe high-vowel harmony, in sentences with expressed non-human plural objects.

A review of some 6,000 example sentences written by three of the native-speaking co-authors of the Marshallese-English Dictionary (Abo et al. 1976) reveals 75 sentences in which this clash and ensuing variation can be seen. Of these sentences with expressed non-human plural objects, 31 used an F-1 or F-2 form that did not agree with the object, and 44 used an F-5 form that did.

| Classes | I | III | V | /kẹn/ | Total |
|---|---|---|---|---|---|
| /ey/ | 3 | 2 | 23 | 3 | 31 |
| /iy/ | 1 | 5 | 38 | 0 | 44 |
| Total: | 4 | 7 | 61 | 3 | 75 |

The preponderance of these examples are in Classes III and V, for which /ey/ is obligatory in F-1 (and therefore in F-2), and for which /iy/ in F-5 is thus replacive. In Class I and for the preposition /kẹn/, the /ey/ is an F-2 extension whose use was optional. The low incidence of examples of this latter type reflects a proclivity for omitting (at least in writing) the optional extensions of F-2 for those classes where that is possible. The same is true for the optional use of /iy/ in II and VII.

Although the above figures would indicate some preference for showing agreement, it is not readily apparent that this is the direction in which the variation may eventually be resolved. If it is true (as will tentatively be concluded) that the antecedent of /iy/ in F-5 was once obligatory, the present situation constitutes a deterioration of the former domain of /iy/ brought about by two factors: (1) the dilution of its role as plural marker by its extension to the singular function F-3 with the rise of high-vowel harmony, and (2) the dilution of its role as object marker with its extension to F-2 as general transitive marker, as part of the tendency toward stem extension.

## THE ORIGINS OF /ey/ AND /iy/ IN THEIR PRESENT FUNCTIONS

Speculations as to origins have accompanied the discussion of some of the individual classes, and it has sometimes been noted how class membership might have changed. Each class of verbs no doubt contains renegade members. The basis for their establishment is a synchronic one: how their transitivity is marked in comparison with the most common extant reminder of their roots, the infinitive counterparts. Because the latter often acquired their own differential markers through reduplication, the burden on transitive marking was lightened, and some transitive markers became optional. This enabled some original members of Classes V and VI, as their /ey/ and /iy/ markers in F-1 became optional, to merge with members of Classes I and II respectively in all functions, and this was the origin of most zero thematic members of these latter classes.

It seems clear that originally Classes I and II were consonant-final, Classes III through VI were vowel-final, and that Class VII (as already noted) may have had either or both origins. Of the originally vowel-final classes, III, IV, and VI can be seen as subtypes of a more general V.

Harrison (1977) has reconstructed three transitive markers for Proto-Micronesian (PMC), and has traced their development in the various daughter languages based on information available. Since all three were vocalic suffixes, and many of the languages have undergone a general deletion of final vowels, most weight has had to be put on evidence from the one language that has escaped this process, KIR. The suffixes and functions posited for PMC are:

-*i, used with pronominal objects,
-*a, used with singular noun phrase objects, and
-*ii, used with plural noun phrase objects and, anaphorically for plural inanimate objects.

Table 2: Markers predicted for Marshallese by positing Proto-Micronesian markers in one possible arrangement with present functions

| FUNCTIONS | PROTO MARKERS | PROTO-MICRONESIAN ROOT TYPES | | |
|---|---|---|---|---|
| | | Consonant-final | High-vowel-final | Non-high-vowel-final |
| F-1 | *-a | ...*C-a > ...C<u>Ø</u> | ...*C[i]-a > ...C<u>ey</u> ∿ ...C<u>Ø</u> | ...*Ce-a > ...C<u>ey</u> |
| F-3 | *-i-a | ...*C-i-a > ...C<u>ey</u> | ...*C[i]-i-a > ...Ciyey ∿ ...C<u>ey</u> | ...*Ce-(i)-a > ...Ceyey ∿ ...C<u>ey</u> |
| F-5 | *-ii | ...*C-ii > ...C<u>iy</u> | ...*C[i]-ii > ...Ciyiy ∿ ...C<u>iy</u> | ...*Ce-(i)i > ...Ceyiy ∿ Cey |

KEY: [ ] enclose high vowels whose reflexes have become optional in Kiribati through paradigmatic pressures that have caused partial reanalysis of high-vowel-final roots as being consonant-final

enclose other high vowels whose reflexes in Kiribati have dropped as a result of phonological processes

underlining indicates plausible present-day markers

One possible matching of these functions with those of /ey/ and /iy/ in present-day MRS is shown in Table 2. F-3, the 3s anaphoric function, is shown as being filled by the reconstructed 3s object -*a following the -*i transitive marker. Final non-high vowels are represented by e in the table, since many earlier a's are known to have raised to e, especially under the influence of neighbouring i's, and since there are no a's to be found in the present-day markers. Changes known to have taken place generally between PMC and MRS are applied to these posited proto verb endings. They include the raising of a to e after i, the development of ę from i and e sequences (in either order) when the second is dropped by final vowel deletion (as reflected in the two alternations involving /ę/ already noted), the 'high-mid harmony' rule, final vowel deletion, and epenthetic semiconsonant insertion. Both longer and shorter proto-forms have been developed for most vowel-final roots (following up on the discovery by Harrison (1977:50-51) of evidence for possible paradigm reanalysis), giving the variants shown in Table 2. In no case (with one possible exception, to be discussed) do the longer forms yield present-day markers. Plausible markers (including /ey/, /iy/, and zero) are underlined singly in the table.

Several interesting points emerge from Table 2. There is a scarcity of present-day markers developing from *-i-a as filler of F-3. In too many cases the predicted development yields a form that is too long, or that contains /ę/ instead of /e/. Only the shorter variant for non-high-vowel-final roots, yielding the /ey/ to be found in the F-3 forms of three of the classes, conforms at all to expectations, but this development is identical to that for the *-a filler of F-1. Thus it can be concluded that the /ey/ filling F-3 today did not develop from *-i-a, but was instead an extension to anaphoric contexts of the *-a filling F-1. This yields plausible results across the board, with zero for consonant-final roots and the high-vowel-final roots reanalysed as such (in the short form developments), and /ey/ for non-high-vowel-final roots. It is thus questionable whether the *-a reconstructed as the 3s object pronoun had achieved suffixal status in PMC, as hypothesised by Harrison (1977:16, 32). If it had, its failure to show the expected reflexes with the various root types might be explained either by its being restricted to animate referents like the other pronouns and thus unavailable for inanimate, or by its reflexes having been replaced by those of the *-a from F-1 at some point.

The developments predicted in Table 2 for the *-a in F-1 for roots with final high vowels are especially interesting, with the longer form yielding an /ęy/ marker (with high-mid vowel), and the shorter ones yielding zero. A scanning of the lexicon for transitive verbs with high-mid vowels (and whose /ey/ markers would be expected to raise to the same height by high-mid harmony) produces some interesting examples. There are some which are still members of Class V, which have a mandatory /ęy/ suffix in F-1 even though their infinitives show reduplication:

| | | | |
|---|---|---|---|
| /jędęy/ | *observe s.* | /jędjęd/ | *observe* |
| /lęqęy/ | *wash oneself* | /lęqlęq/ | *wash female genitals* |
| /tępęy/ | *give gifts to s.* | /tęptęp/ | *ritual gift-giving* |

There are others whose /ęy/ suffix has become optional, thus shifting them to zero thematic status in Class I. One of these — /jęm/ *sharpen s.* — has already been cited in the discussion of that class. Another is

| | | | |
|---|---|---|---|
| /lęt/ | *wear s.* | /lętlęt/ | *wear for the first time* |

Most interesting is the infinitive /jękjęk/ *chop* which has variant transitive stems /jek/ (Class I) and /jękęy/ (Class V) *chop s.*. These are precisely the variant developments that would be posited for *-a in F-1 with a proto-root *jeki.

| Long Form | Short Form | Developments |
|---|---|---|
| *jekia | *jeka | |
| jekie | — | (a > e /i__) |
| jekęe | — | (i > e /__e#) |
| jekę | jek | (final vowel deletion) |
| jękę | — | (high-mid harmony) |
| jękęy | — | (epenthetic semiconsonant) |
| /jękęy/ | /jek/ | |

The verbs with high-mid vowels previously cited have undergone only the long form development, or have lost the historic short form. A verb like the one below has undergone only the short form development, or has lost the long form:

/tel/ *put s. inside* /tęltęl/ *collect*

Note that verbs like /jek/ and /tel/ now take /ey/ in F-2, rather than the /ęy/ they should have gotten historically through the long form development. The discovery of verbs of this type strengthens the assertion made above that *-a in F-1 produces plausible results across the board.

Another item of interest from Table 2 is the lack of present-day markers in the development of *-ii in F-4 and F-5 for roots in final non-high vowels. It must be concluded that such roots received their modern /iy/ marker in these functions by analogy with the other types of roots, which show this development normally. Even more striking is the lack of development of any /iy/ marker in either of the singular object functions F-1 or F-3 when only *-a and *-i-a are posited as sources. Such development, if it occurred, might be a means of explaining the origin of the classes with high vowels that neutralise markers (Classes II and VI). It should be noted that even the zeros that develop from *-a in F-1 would have as concomitant development a high-mid vowel preceding the stem consonant, from an earlier high vowel as part of the iCe- / eC# alternation. Such roots, originally with high vowel preceding their final consonant, with that vowel thus converted to high-mid, would hardly form an inviting base for analogising the /iy/ marker from its plural functions in these classes. They would have been more likely to attract the /ey/ from where it is thought to have developed in F-1, with the final-non-high-vowel roots, and raise it by the high-mid harmonising rule.

Thus it seems necessary to conclude that *-a did not fill F-1 for original members of Class II, those whose roots were consonant-final with a preceding high vowel, but that *-ii did instead. It should be noted that Harrison (1977: 10-12) found it necessary to posit the same deviation for all consonant-final roots in Trukic, not just those with preceding high vowel. This leads to the conclusion that the distribution of the reflexes of the *-a, *-i, and *-ii markers in KIR should not be reconstructed as firmly fixed in PMC, since both Trukic and MRS evidence points with partial agreement to a partially different distribution. There almost certainly were dialectal differences in PMC. If, as hypothesised, these markers were primarily transitive markers and only secondarily object markers depending on their selection with different types of objects, it should not be too surprising if they came to be sorted out differently in the various dialects, subject to phonological influences, among others.[12]

Parallel use of the *-ii marker should not be posited in F-1 for all final-high-vowel roots, for as has been noted, the positing of *-a in this function gives felicitous results for verbs such as /jek/ ∿ /jękęy/ *chop s.* < *jek(i)-a.

Table 3: Summary of possible account of the origins of present fillers of F-1, F-2, F-5, and 'F-6' for original members of the various transitive verb classes.

| Functions | Proto-Suffixes | VERB CLASSES AND CANONICAL FORMS OF ORIGINAL ROOTS BEFORE SUFFIXATION (WHERE e STANDS FOR NON-HIGH V) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | V | IVa | IVb | V | III | I | I | II | VII | VI |
| | | jękęy<br>jękjęk | jey<br>jęyjęy | ley-DIR<br>ley-DIR | beqey<br>beq | biqey<br>biqabęq | jek<br>jękjęk | deget<br>degdeg | m̧ijit<br>m̧ijm̧ij | biwik<br>biw | tir̈iy<br>tir̈ |
| | | ...eC(i)<br>long | ...ey(i) | ...eCe-DIR | ...eCe | ...iCe | ...eC(i)<br>short | ...eC | ...iC | ...ik | ...iC(i) |
| F-1 | *-a | -ęy | ø | ø | ey | ey | ø | ø → | ø → | ø | iy |
| F-2 | *-a | [ęC]ęy | [ey]ø | [e] ey ← | ey | ey → | ey → | ey | | | iy |
| | *-ii | | | | (ęy) | (ęy) | | | iy | iy | iy |
| F-5 | *-ii | iy | [ęy]iy | [i]iy → | iy | iy ← | iy | iy | iy | iy | iy |
| 'F-6' | *-i | (iy) | ([ęy]iy) | ø | (ęy) | (ęy) | (ęC) | (ęC) | ø | ø | ø |

Note: Predicted historical development of each suffixed root type is underlined, and listed by its present-day function. Functions filled by extension from other functions, or by analogy with other classes, are shown by arrows pointing from the supposed source. Predicted developments thought to have been replaced are enclosed in parentheses. Square brackets enclose present-day stem material of special interest.

The *-ii marker in F-1 can be posited, however, for final-high-vowel roots that also had a high vowel preceding the last consonant — those in ...iCi, or at least those in ...iCi that were reanalysed as consonant-final, and could then be expected to undergo a short-form development paralleling that outlined above for original members of Class II in singular functions. This then provides an explanation for original members of Class VI: with final vowel deletion, the remaining /i/ reflex of the *-ii marker (and epenthetic /y/) became the distinctive transitive marker in F-1, the only formal difference between such verbs and their infinitives — where the latter had not reduplicated. This was not the case for original members of Class II, which had thematic material available to make the distinction, and could leave the /iy/ optional, at least in non-anaphoric contexts. The positing of ...iCi roots (which like original members of Class II selected the *-ii marker for singular functions as well as plural) as original members of Class VI makes sense from a synchronic point of view, since the final i of such roots would have been lost in nonsuffixed forms such as infinitives, and since today it is the reflex of the original penultimate i that governs the selection of the /iy/ marker, and the neutralisation of functions that it entails.

The predictions of Table 2, modified according to the preceding discussion, are summarised in Table 3, as constituting a possible account of the origins of the /ey/ and /iy/ formatives in their contemporary functions. In Table 3, the order of the classes has been rearranged from their numbered sequence (as in Table 1) to better display the interclass influences posited. As can be seen, the /ey/ formative began as the root-final non-high vowel of original members of Class III and V. When this vowel was lost from unsuffixed forms and the *-a suffix was also deleted from transitive stems by final vowel deletion, it took over the distinctive transitive marking function of the latter. It was then available for extension in the optional F-2 to Class I stems, whose transitivity was already sufficiently marked in F-1 by thematic material. On the other hand, the /iy/ formative is the continuation of an original suffix, which with original double or long vowel, was not totally obliterated by final vowel deletion. The /ey/ that developed from the *-ii suffix in the plural functions of Classes III and V was levelled by the /iy/ that developed elsewhere. It is ironic that this levelling should have occurred in precisely the same classes that were the sole source for /ey/ in the singular.

Little can be said with certainty concerning the *-i reconstructed as transitive marker before pronouns, other than that if it ever functioned in that capacity for MRS, it has disappeared without a trace. It should have left traces of two types other than zero, as indicated for 'F-6' in Table 3: roots ending in i should have suffixal /iy/ in their longer developments, and roots ending in e or eC should have had those mid vowels raised to high-mid as part of the ...eCi- / ęC# alternation.

This latter development may have taken place but been reinterpreted as the result of a synchronic rule, the high-mid harmony rule, because of the coincidence that all the object pronouns have high-mid vowels. The rule not only can be viewed as a constraint on morpheme structure, requiring that mid and high-mid vowels not co-occur, but also operates as a low-level phonetic rule, raising both prefix and suffix mid vowels with high mid stems. The latter often extends to elements within compounds and phonological phrases, usually regressively, so that the second syllable of the compound /jiten-bęręw/ *sweetheart* may be pronounced with [ę] while being felt to have underlying /e/, and — most relevant here — the second syllable of verbs such as /piqet/ *look for s.* may be pronounced with [ę] before object pronouns, again while maintaining underlying /e/:

11. / yi-har piqet yęq /.
lsSM-PAST *look for* 2sOBJ

*I was looking for you.*

Thus, as in this example, if transitive verbs such as /piqet/ originally took an *-i marker before pronoun objects such as /yęq/, when this marker was lost by final vowel deletion and the preceding /e/ concomitantly raised to /ę/, such raising could have been attributed instead to the high-mid harmony rule operating in an environment in which the next morpheme always had a high-mid vowel. If this is what happened, the other vestiges of the *-i transitive marker — the /iy/ developed for roots ending in i — must have been dropped through paradigm levelling.

## NOTES

1. This paper owes its origin to the stimulating and insightful presentation of S.P. Harrison (1977) to the Second International Conference on Austronesian Linguistics held in Canberra in January 1978. I am indebted to Robert Blust, Alfred Capelle, Michael Forman, Sheldon Harrison, Paul Hopper, and Laurence Thompson for comments made on an earlier version.

2. The orthography is that of Abo et al. (1976). The phonemic transcription from the same work is used to show morphological detail, and may be summarised briefly as follows: four vowels differing among themselves only along the height dimension (/i/ high; /ę/ high-mid; /e/ mid; /a/ low) are front next to unmarked (phonetically palatalised) consonants (/y/; /p/, /j/ [tʸ]; /m/, /n/; /l/, /d/ [rʸ]), back (but unrounded) next to pharyngealised or velarised consonants (/h/; /b/, /t/, /k/; /m̦/, /n̦/, /g/ [ŋ]; /l̦/, /r/), and backed and rounded next to rounded consonants (/w/; /q/ [kʷ]; /m̈/, /g̈/ [ŋʷ]; /l̈/, /r̈/). Vowels between consonants from different of these three sets, as a result of the counter influences of the consonants, may be perceived as breaking, or as of an intermediate quality. Excrescent vowels of predictable quality are produced between contiguous consonants that are not homorganic.

3. The basic demonstrative system is as follows:

| 'Person' | Singular | | Pl. Non-human | | Pl. Human | |
|---|---|---|---|---|---|---|
| '0' | eo | /yew/ | ko | /kew/ | ro | /rew/ |
| 1e | e | /yęy/ | kā | /kay/ | rā | /ray/ |
| 1i | in | /yin/ | kein | /kęyin/ | rein | /ręyin/ |
| 2 | n̦e | /n̦ey/ | kan̦e | /kan̦ey/ | ran̦e | /ran̦ey/ |
| 3 | en̦ | /yen̦/ | kan̦ | /kan̦/ | ran̦ | /ran̦/ |
| '?' | ewi | /yęwiy/ | erki | /yerkiy/ | erri | /yerriy/ |

The person labels traditionally associated with pronouns seem apt for extension to the basic semantic distinctions among the rows of the system, with even the inclusive-exclusive distinction being observed in the first person. Thus, kā, being 'first person exclusive plural non-human', might be translated as *these things (closer to me than to you)*. The 'zero-person' demonstratives are used for definite referents not in view, and also serve

as relative pronouns. The last row (labelled with a question mark) contains locative interrogatives, which are formally similar to the demonstratives, and which occur in some of the example sentences used in this paper. It may be noted that this system counters a putative language universal by maintaining in the plural a contrast — human-non-human — not found in the singular.

4. Reconstructed and contemporary transcriptions are intermingled in the following example for simplicity of exposition.

5. The vowel alternation exemplified here has been termed "low vowel dissimilation" (Bender 1969).

6. The vowel alternation exemplified here is discussed later in the paper, and in Note 11.

7. Ward Goodenough (1963) has demonstrated the origins of initial double consonants in Micronesian languages in the reduplication of initial syllables, with later vowel loss.

8. Such sentences are dismissed by adult speakers as possible only on the part of children learning the language, or of non-native speakers.

9. As found, for example, in Grace 1969, and subsequent unpublished updatings of this list.

10. Only the low vowel /a/ escapes syncope in reduplicated forms such as this.

11. The two alternations seem to serve the purpose of reducing the homophony that would otherwise result from final vowel deletion, by creating a new intermediate quality in the penultimate vowel that remains. An umlaut explanation is unsatisfactory, since assimilation does not generally take place when the final vowel is preserved by suffixation. Only those final vowels that are lost leave their imprint on the preceding vowel: final vowels that are mid when the preceding vowel is high, or vice versa.

12. Jeff Marck (personal communication) has recently discovered a semantic contrast between stems formed with i and a transitivising increments on one and the same root in Saipan Carolinian, with i-stems having a locative and/or partitive meaning, and a-stems a more perfective and wholly transitive meaning. Parallel phenomena have since been attested in most other Trukic languages by Frederick Jackson (personal communication). This phenomenon seems limited to the Trukic group within Micronesian.

## BIBLIOGRAPHY

ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DeBRUM

1976 *Marshallese-English dictionary.* Honolulu: University Press of Hawaii.

BENDER, Byron W.

1969 Vowel dissimilation in Marshallese. *Working Papers in Linguistics, University of Hawaii* 1/1:88-96.

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The proceedings of the Ninth Pacific Science Congress, 1957*, vol.3:77-86.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

HARRISON, Sheldon P.

1978 Transitive marking in Micronesian languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 1067-1127. *Pacific Linguistics*, C-61.

Bender, B.W. "Object marking in Marshallese". In Bender, B.W. editor, *Studies in Micronesian linguistics.*
**C-80:443-465**. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.443*

# SOME SYNTACTIC PROCESSES IN KIRIBATI

Roderick A. Jacobs

## I. INTRODUCTION

I describe here some major syntactic processes in Kiribati (Gilbertese) (KIR) and consider their relevance to general syntactic theory and to problems in the description of the syntax of Oceanic languages. The general theoretical framework assumed here is a semantically-oriented modification of Chomsky's so-called 'Standard' theory (Chomsky 1967), one similar to that used in the present writer's comparative syntactic study of three Amerindian languages (Jacobs 1975). The data discussed here are drawn from texts and from elicitation sessions with KIR speakers, both in Hawaii and in Kiribati (formerly Gilbert Islands). Two brief descriptions exist — Bingham 1861 and Cowell 1954, a series of grammar lessions and exercises for non-native speakers. Sohn 1973 includes a useful discussion of relativisation in KIR.

## II. EQUATIVE AND PREDICATIVE SENTENCES

Sentences in KIR fall into two major classes: equative and predicative. Equative sentences are verbless sentences in which both the subject and the predicate phrase are noun phrases. The subject noun phrase typically expresses 'given' information, information assumed to be familiar already to the addressee while the predicate noun phrase is the major assertion of the clause. What is asserted may be a role predication, as in (1) below, or an identity predication as in (2):

1. Nakaa (bon) te tia-moti.
   ASSERTIVE ART *judge*

   *Nakaa is/was (a) judge.*

2. Te tia-moti arei (bon) Nakaa.
   ART *judge* *that* ASSERT

   *That judge is Nakaa.*

Equative sentences have no aspectual marking. We will discuss some apparent exceptions to this claim later. The optional assertive particle bon introduces the predicate phrase, the main assertion of the sentence.

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 467-490. *Pacific Linguistics*, C-80, 1984.

The linear ordering **SUBJECT-**(bon)**-PREDICATE PHRASE** of (1) and (2) above seems less common than the ordering (bon)**-PREDICATE PHRASE-SUBJECT**, as in

3. Bon tama-u Nakibwae
ASSERT *father-my*

*Nakibwae is my father.*

With such an ordering the assertive bon is almost never omitted, suggesting that this may be a more marked ordering for equative sentences.

Predicative sentences consist minimally of a subject proform, which appears to be prefixed to the predicate phrase, and the predicate phrase consisting minimally of a main verb. The obligatory occurrence of the subject proform is a major distinguishing criterion for predicative sentences. Equative clauses have no subject proforms. Following the predicate phrase is an optional noun phrase, the referential 'antecedent' of the subject proform. Where the information represented by this noun phrase is quite clear from the context, it is frequently omitted. Of course the obligatory subject proform marks the role of this noun phrase, as in

4. A-maeka ikai.
*they-live here*

*They live here.*

although, as mentioned above, the a *they* can be more fully specified:

5. A-maeka ikai ataei-n-aine akekei
*they-live here young-of-female those*

*Those girls live here.*

The predicate phrase may also include negative and aspectual markers, direct and indirect object noun phrases, and prepositional phrases. The KIR equivalents of English predicate adjectives may be considered to be intransitive verbs.

Lyons (1978) has argued that it is valid, even necessary, for the investigator of the grammar of a language to seek to determine how and to what extent the grammatical structure of a sentence determines its meaning regardless of situational factors. These include contextual appropriateness, speech act function, politeness rules, and, more obviously, such performance factors as stammering, slips of the tongue, and even differences arising from minor dialectal and idiolectal variations, and age and sex differences. Such a heuristic may involve delicate decisions by the investigator but it appears essential for effective investigation of the syntax of a language. In the case of KIR, sentences like (4) above are probably more common than sentences like (5) with its noun phrase specification, ataei-n-aine akekei *those girls*, of the subject proform a *they*. But we consider the more fully specified ones more basic to an understanding of the system of the language. The subject proform can be considered a kind of grammatical agreement marker, one which serves as a 'trace' if the noun phrase it stands for is omitted.

For equative sentences we noted earlier an indication that the ordering **SUBJECT-**(bon)**-PREDICATE PHRASE** was the less marked ordering. But unless we identify the subject proform as a noun phrase, the subject of its clause, predicative sentences in KIR seem to be predicate-initial as they actually occur. We assume a framework in which there is a linearly unordered semantic deep structure with hierarchical structuring to capture semantic scope relations. At some fairly deep level, linearisation occurs according to

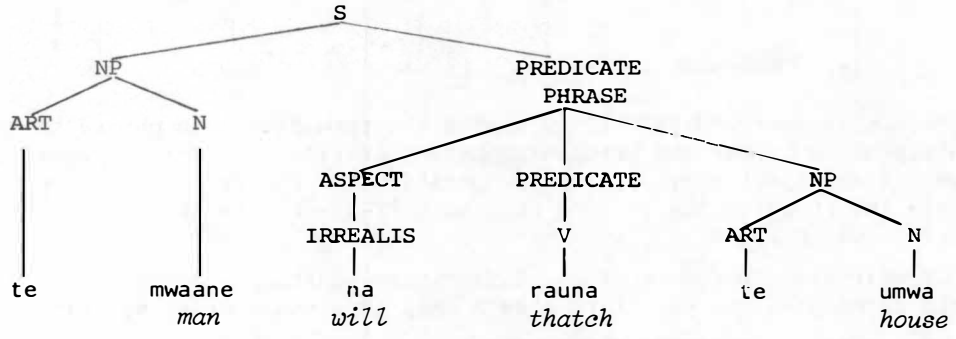
language-specific restrictions and the informational status of the content. For the predicative sentence (6):

6. E-na rauna te umwa te mmwaane
*he*-IRREALIS *thatch* ART *house* ART *man*

*The man will thatch the roof.*

we posit an earlier state of the derivation which can, with details omitted that are nor immediately relevant, be shown thus:

7.




with a yet earlier stage marking topic-comment structure,[1] much as in the so-called topicalised forms:

8. Te mmwaane, e-na rauna te umwa.

*As for the man, he'll thatch the house.*

9. Te umwa , e-na rauna te mmwaane
ART *house, he-will thatch-it* ART *man*

*As for the house, the man will thatch it.*

10. Te umwa, e-na rauna-aki iroun te mmwaane.
*house it-will thatch*-PASSIVE *by* ART *man*

*As for the house, it will be thatched by the man.*

A structure like (7) for sentence (6) suggests that some sort of agreement phenomenon produces the subject copy, especially since other nuclear Micronesian (MC) languages have surface structures corresponding either to (7) or to (7) with a subject proform at the beginning of the verb phrase, e.g.

11. Ohl-o mwemeitla Ruk
*man-that visit Truk*

*That man visited Truk.*

(Ponapean)

12. Mwal-la e-pwe la
*man-that he-will go*

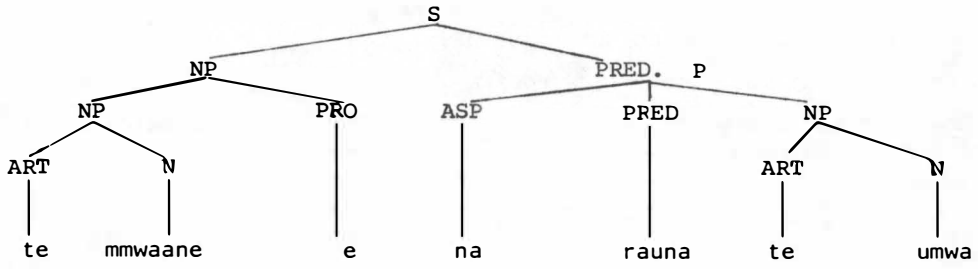
*That man will go.*

(Saipan Carolinian)

Moreover, a fairly similar marking phenomenon has an object suffix attached to the verb if the object does not immediately follow the verb and is animate. Full noun phrases seem to determine affix proforms rather than vice-versa.

To convert (7) into sentence (6), then, a SUBJECT COPYING rules adds a pronominal copy, possibly by Chomsky-adjunction as in (13):

13.




and then obligatory EXTRAPOSITION shifts the specified noun phrase te mmwaane to the position after the predicate phrase, leaving the subject proform (henceforward the subject copy) in initial position. A fairly late morphophonemic process incorporates the subject copy as a prefix to the first constituent of the predicate phrase.

Predicative sentences, then, differ from equational sentences in having a verb as head of the predicate phrase and, in surface ordering, having

14. $COPY_i$ - PREDICATE PHRASE ($NP_i$)

(where subscript i marks assumed coreference) while equatives have two common orderings:

15. NP - PREDICATE PHRASE

or

16. PREDICATE PHRASE - NP

A further difference is that equative sentences allow no aspectual marker:

17. *Nakaa na te tia-moti
IRREALIS ART *judge*

*Nakaa will be the judge.*

18. *Nakaa a te tia-moti
REALIS ART *judge*

*Nakaa is/was the judge.*

However the subject may be one of the set of independent (or 'emphatic') pronouns:

19. Bon te tia-moti ngaia
ASSERT ART *judge* *he*

*He is the judge.*

20. Ngaira bon taani-moti
*we* ASSERT *judges*

*We are the judges.*

More common still are sentences in which these pronouns are predicate noun phrases.

21. Boni ngaia te tia-moti
ASSERT *he* ART *judge*

*It's HIM that's the judge.*

22. Boni ngaira taani-moti
ASSERT *we* *judges*

*WE are the judges.*

A final difference is that predicative sentences are negated with aki:

23. E-aki atai-a
*he not know-him*

*He doesn't know him.*

while equative sentence have tiaki:

24. Tiaki te tia-moti Nakaa
*not* ART *judge*

*Nakaa is not the judge.*

## III. TOPICALISATION AND ASSERTIONAL STATUS

In the previous section EXTRAPOSITION was described as an obligatory transformation. Sentences like (6), repeated here as (25), represent the usual constituent ordering for predicative sentences:

25. E-na rauna te umwa te mnwaame.
*he*-IRREALIS *thatch* ART *house* ART *man*

*The man will thatch the house.*

But there also exist sentences in which a full noun phrase subject has apparently not been extraposed — as in (8), repeated here as (26):

26. Te mmwaane, e-na rauna te umwa
ART *man* *he*-IRREALIS *thatch* ART *house*

*As for the man, he'll thatch the house.*

But here an intonational break, shown above with a comma, indicates that te mmwaane is not an ordinary subject noun phrase but the topic noun phrase.

It might seem reasonable to generate both (25) and (26) from a single underlying structure, say (7) above. After all, the two sentences are cognitively synonymous in the sense that the same truth conditions hold for both. After SUBJECT COPYING, an optional TOPICALISATION rule could apply, raising a noun phrase (not necessarily the subject) into topic position. Finally, if the subject is not topicalised, the EXTRAPOSITION rule must apply.

But such a solution seems unsatisfactory since it implies that (25) and (26) are exact paraphrases. Like their English translations, (25) and (26) can be almost exact paraphrases. But the environments in which (26) may be uttered felicitously are a special subset of those for (25), which is not atypical of forms elicited as isolated sentences. In fact (26) is part of a longer discourse, a small segment of which is given below as (27). Here the boy Toaa is talking to his older sister, Nei Ribwa, after a storm had destroyed their house:

27. Toaa: Ao taraia bwa iai Nakibwae ao Bwatuku —
*And look! that there's Nakib'ae and Bwatuku —*

ao te mmwaane! I-aki atai-a. Bukin teraa bwa a-roko ikai
*and a man! I-not know-him. Why that they-come here?*

Nei Ribwa: Nakibwae ao Bwatuku, a-roko ni ibuobuoki —
*and they-come to help*

ni urakinii kaai aika kaina. Te mmwaane, e-na
*to bring wood which pandanus. The man, he-will*

rauna te umwa.
*thatch the house.*

Sentence (26) is spoken after the man in question has already been referred to. In an interesting sense, the topicalised noun phrase functions in the discourse much as a relative pronoun might function in a single sentence. Both relate a proposition to a previously mentioned referential noun phrase. The major informational difference is that, when (26) is uttered, the clause e na rauna te umwa, *he will thatch the roof* is still new information — an assertion. A little later, when the man is gone, Toaa refers to him thus when his mother asks what has happened:

28. E-a roko mai te unimmwaane ae e-na rauna te umwa.
*He*-REALIS *come here* ART *old-man that he*-IRREALIS *thatch* ART *house.*

*An old man who's going to thatch the house came here.*

The relative pronoun ae reduces the informational prominence of the clause it introduces, so that the clause is part of a referential noun phrase rather than the major assertion of the sentence. It must be noted however that the relative clause here still retains a degree of assertional prominence and communicates information believed by the speaker to be new to the hearer.

It is obvious that one major reason that generative grammars have treated topicalised noun phrases so differently from relativised noun phrases is that the former present major difficulties in formalisation for a sentence grammar. A single set of interrelated semantic propositions can take on a variety of syntactic and morphological forms. The forms are in part determined by supra-sentential speaker-hearer relations. Assumptions as to the style and content of a discourse are subject to revision as the discourse is under way. The participants continually structure and restructure the content. Asserted ('new') information in one sentence becomes, by regular processes, assumed ('given') information later in the same sentence or in a subsequent sentence. The simple assertion (with a passive verb):

29. Ko-a tangir-aki irou-u.
*you*-REALIS *love*-PASV *by-me*

*You are beloved of me.*

may in a subsequent sentence be represented as a referential entity, a nominalisation:

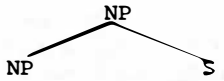
30. tangira-m irou-u
*(the)loving-you by-me*

*my love for you*

Nominal structures are typically used to refer rather than assert. Suprasentential discourse factors are the major determinants governing the occurrence of forms like (30) rather than (29).

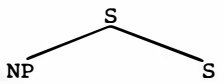
The topic noun phrase in (26), te mmwaane, *the man*, is a resumptive element referring back to something earlier in the discourse or otherwise assumed to be known to the addressee. In such sentences the subject copy e-, if there is no other noun phrase in the clause that it refers to, refers back either to the topic or to some other already familiar entity. In sentences in which a full noun phrase follows the predicate phrase as its extraposed subject, the e- refers forward. Where the subject copy refers back to the topic noun phrase, the relation between the topic and the following predicate phrase is not unlike that holding between a head noun phrase and its relative clause, which in the generative tradition is usually shown as (31):

31.




though we have argued for a somewhat different structure in the paper on Trukic languages elsewhere in this volume. If the corresponding structure for a topicalised sentence is

32.




then the fact that the NP S in (31) is dominated by an NP rather than an S, as in (32), corresponds rather neatly to the difference in informational status between relative clauses and topicalised sentences.

But both forms represent points along a continuum of assertive prominence. The topic noun phraase, despite its status as given information, is more prominent than the subject noun phrase in the untopicalised (25). A noun phrase could, of course, be yet more prominent if it were treated as new information, i.e. as a predicate.

Indeed it is possible for a noun phrase to constitute an independent sentence:

33. Te mmwaane.
ART *man*

*It's a man.*

An Englishman, seeing a man with a long white wig strolling towards the Old Bailey court buildings in London, might remark to a non-British friend:

34. A judge.

But we might prefer to mark the predicative status of the noun phrase more clearly:

35. It's a judge.

The availability, indeed the preferability, of bon in Gilbertese for (36):

36. (Bon) te mmwaane. E-na rauna te umwa.

*It's a man. He's going to thatch the house.*

indicates that te mmwaane is actually a predicate noun phrase here.

Another option determined, in part at least, by discourse function is a sentence in which te mmwaane and e-na rauna te umwa reverse their normal informational roles. The noun phrase te mmwaane, as we have already seen, may serve as predicate phrase. In (37) it is predicated of the clause e-na rauna te umwa.

37. Bon te MMWAANE are e-na rauna te umwa.
ASSERT ART *man* *which he*-IRREALIS *thatch* ART *house*

*It's the MAN who's gonna thatch the house.*

In KIR, as in Eng., it is possible to have a different version in which the primary sentence stress is on the last word:

38. Bon te mmwaane are e-na rauna te UMWA.

*It's the man who's gonna thatch the HOUSE.*

In (37) the predicate phrase is just te mmwaane, while in (38) the predicate phrase includes everything after bon. We claim that are e na rauna te umwa in (37) is not part of the predicate phrase but rather an extraposed noun phrase which was the underlying subject. So (37) is derived from an underlying structure something like that of (39):

39. Are e-na rauna te umwa bon te MMWAANE.

**Who will thatch the house is the MAN.*
*The one who will thatch the house is the MAN.*

The subject noun phrase in (39) consists of a clause introduced by a relative proform are, i.e. the subject is a relative clause without a noun phrase head. The KIR version is perfectly acceptable although its Eng. counterpart requires a NP head like *the one*. Such headless relative clauses are not uncommon in texts but are a little less frequent in modern spoken KIR. I suspect the discrepancy has resulted from the impact of Eng. Of course in Eng. too such clauses were formerly more common:

40. Who steales my purse steales trash.

Shakespeare, *Othello* III, 3.157.

41. Whom the gods love die young.

Byron, *Don Juan*, 4.12.

Now, if (39) indeed has a relative clause as its subject NP, then (39) consists of two major constituents: a noun phrase subject and a noun phrase serving as predicate. So (39) is an EQUATIONAL SENTENCE relating two noun phrases to each other. And the extraposed version (37) is likewise an equational sentence.

But sentences like (37) present a problem. The predicate phrase is the noun phrase te mmwaane are e na rauna te umwa. But (37) has no overt subject noun phrase. Nor has sentence (42):

42. (Bon) te tia-moti
ASSERT ART *judge*

*It/he is the/a judge.*

Yet it is clear that the predicate noun phrases are being posited of some entity considered as 'given' and that underlyingly (37) and (42) are equative sentences. Some speakers in fact use a version of (42) with a subject proform:

43. E -bon te tia-moti.
*it/he*-ASSERT ART *judge*

*It/he-is the/a judge.*

Informants describe sentences like (43) as most often produced by KIR children. But (43) is much less common if the assertive bon is omitted. On the other hand, with an additional aspectual element intervening, many adult speakers accept the sentence:

44. E-na bon te tia-moti
*he*-IRREALIS ASSERT ART *judge*

*He'll be the judge.*

Subject copies are unacceptable with predicate noun phrases in verbless sentences. But apparently the intervention of elements that normally introduce verbs allows many speakers to 'forget' this restriction. Since -na cannot introduce a predicate noun phrase, the bon is the true culprit causing the speaker to 'forget'.

45. *E-na te tia-moti Nakaa
*he*-IRREALIS ART *judge*

*Nakaa will be the judge.*

However a very similar form is perfectly acceptable:

46. E-na tia-moti Nakaa
*he*-IRREALIS *judge*

*Nakaa will be the judge*

Sentence (46) means that Nakaa will judge, i.e. act as judge, take on the role of judge. Likewise a sentence like (47):

47. E-na George Washington.

means that someone is going to play the role of George Washington in a theatrical performance. The forms tia-moti and George Washington have become verblike enough in (46) and (47) to forbid the occurrence of the article te. Both (46) and (47) are predicative sentences and therefore have subject copies. This also accounts for the unacceptability of tiaki negation and the acceptability of aki negation:

48. *E-na tiaki tia-moti.

49. E-na aki tia-moti

*He won't be the judge.*

These facts might be represented in the grammar by a surface filtering restriction, one which blocks subject copies or aspectuals before noun phrases, with some idiolectal variation where bon intervenes. Or we might avoid such a device by generating subject copies for both predicative and equative sentences and positing a rule deleting the copies before noun phrases. The same idiolectal variation can be incorporated in such a rule. Where an aspectual marker intervenes, predicate noun phrases must be converted into verbs. We might use the term denominal for such ex-NP's. We will, for the purposes of this paper, choose the second alternative and refer to the two transformations needed as SUBJECT COPY DELETION and DENOMINAL FORMATION.

## IV. THE UNDERLYING LIKENESS OF EQUATIVES AND PREDICATIVES

The equative/predicative distinction now appears more like a surface distinction arising out of the lexical category status of the head of the predicate phrase. While predicatives have the surface ordering:

50. $\text{COPY}_i$ — PREDICATE PHRASE — $\text{NP}_i$

where the $\text{NP}_i$ has been extraposed, equative sentences have been described as allowing two orders:

51. PREDICATE PHRASE — NP

or

52. NP — PREDICATE PHRASE

The order (52), which earlier was argued on other grounds to be less marked, may be the basic underlying form, as it may be for predicative sentences. So equative sentences like (53) and (54)

53. Nakaa bon te tia-moti

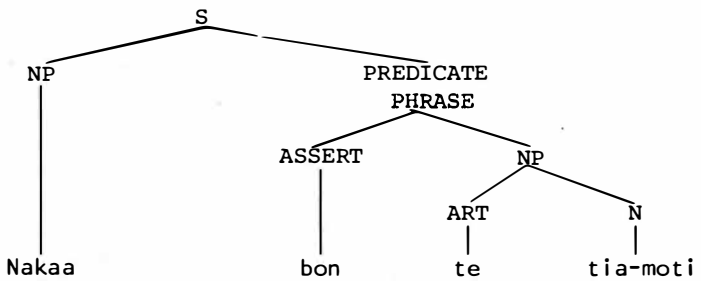
*Nakaa is the judge.*

54. Bon te tia-moti Nakaa

*Nakaa is the judge.*

though very, very slightly different in use, share a common underlying configuration we might show as (55):

55.




To this, as to all other sentences, SUBJECT COPYING applies, yielding an intermediate

56. *Nakaa e-bon te tia-moti

Now EXTRAPOSITION can apply, yielding

57. *E bon te tia-moti Nakaa.

a form like many produced by Kiribati children. But for most adult speakers, whether or not EXTRAPOSITION has applied, COPY DELETION must apply, resulting in (53) or (54) above.

The creation of a subject copy followed by its deletion is not as inelegant as it at first seems, since this deletion applies only to a subset of the cases to which the SUBJECT COPYING rule applies and also allows for the operation of an alternative DENOMINAL FORMATION rule for forms like (46) and (47) above. A more significant question, perhaps, is why EXTRAPOSITION is optional for equatives and obligatory elsewhere. It is noteworthy that where bon occurs in the predicate, extraposition is far more likely. Certainly bon seems to be required in such extraposed forms.

It seems as if, when both subject and predicate are noun phrases, speakers are unclear as to which is the predicate. Hence the wavering over extraposition. This confusion should hardly be surprising since the semantics reinforces it. After all x equals y means also that y equals x. The occurrence of the predicate-marking particle bon clearly identifies the constituent eligible for extraposition, hence the rarity of unextraposed bon sentences. It may be that these rare unextraposed bon forms represent somewhat different discourse assumptions, assumptions which could be represented in a more adequately worked-out version of this framework in terms that could block the application of EXTRAPOSITION.

On this view, then, the distinction between equational sentences and predicative sentences is a fairly superficial one. The surface differences arise from probable processing difficulties caused by the use of the same major category for quite distinct informational functions. In this connection it is interesting to note the existence of strategies like the occasional omission of the article before predicate noun phrases, especially role predicates, in other languages. This creates a possible contrast between subject noun phrases and predicate noun phrases:

58. Jacques est (le/un) professeur.

    *Jacques is a teacher.*

59. You be (the/a) judge, ffotheringay.

60. *Professeur est Jacques.

61. *Judge was ffotheringay.

62. *Vice-President was Agnew.

## V. RELATIVISATION

We have already discussed data indicating that relative clauses in KIR are noun phrases. At first sight relative clause constructions look much like their Eng. counterparts:

63. te mmwaane are ko-noori-a ngkoananoa
    ART *man that you-see-him yesterday*

    *the man that you saw yesterday*

64. te mmwaane are e- ka -mate-aki
    ART *man that he*-CAUSE-*die*-PASV

    *the man that was killed*

65. te kai-n -tekateka are ko-tekateka i- ao -na
    ART *stick-of-sit that you-sit on-surface-its*

    *the chair that you sat on*

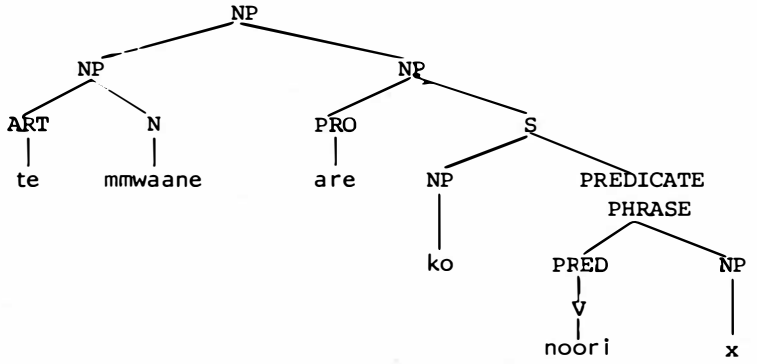
66. te umwa ane ti-nako iai
    ART *house that we-go there*

    *the house that we went to*

Perhaps the most obvious difference is that the relative clause always has a trace proform assumed to be coreferential with the noun phrase head. In (63) the proform is the object suffix, in (64) the subject copy, in (65) the

possessive suffix on the nominal preposition, in (66) the locative pronoun iai. We assume that the head noun phrase forms a single noun phrase along with the relative noun phrase:

67.




The relation between the head noun phrase and the relative noun phrase is the 'presupposed' equivalent of the equative relation of (68) and its unextraposed counterpart (69):

68. Bon te mmwaane are ko-noori-a
ASSERT ART *man* *that you-see-him*

*It's the MAN that you saw.*

69. Are ko-noori-a bon te mmwaane.
*that you-saw-him* ASSERT ART *man*

*The one that you saw was the man.*

Such relative noun phrases can be subjects as in (69) above, and objects as in (70):

70. I-a tia-n noora are ko-uoti-a
*I*-REALIS COMPLETIVE *see* *that you-bring-it*

*I have seen what you brought.*

71. I-a tia-n noora ake ko-uoti-ia
*I*-REALIS COMPLETIVE *see* *those you-bring-them*

*I've seen the things you brought.*

In all these cases the so-called coreferential proform, we claim, is simply an entity specified, like the one/ones, for number and, in KIR, for certain other dimensions marked on the introducing proform (are etc.). It is the construction itself that asserts (in equatives) or presupposes (in relative constructions) the equative relation referred to as coreference.

The proforms 'linking' relative sentences to their heads are quite different from those in other nuclear MC languages. The actual forms are

72.

| | SINGULAR | PLURAL |
|---|---|---|
| 1. | ae | aika |
| 2. | ane | akana |
| 3. | are | ake |

These forms express differences in time, location, or discourse perspective which we need to investigate further. The (1) forms are the here and now forms,

often introducing entities or qualities not referred to before or perhaps generally true. They are the forms most often used in intransitive adjectival verbs:

73. te waa ae (e-) uraura
*the boat that (it) red*

*The red boat*

The parentheses around e- indicate here that the vowel is elided because of the preceding e. The same elision occurs for the plural a-, *they*. However this elision occurs only when the proform would be prefixed to the final word in the clause, normally an intransitive verbal. The forms marked (2) are typically discourse-oriented, introducing material previously referred to in the discourse, often by the addressee. Hence there is also a connection with second-person forms.

The (3) forms often refer to past time outside the discourse and are the most common introducers of longer relative clauses.

Parallel with these proforms are the set of independent demonstratives:

74.

| | SINGULAR | | PLURAL | |
|---|---|---|---|---|
| 1. | aei | *this* | aikai | *these* |
| 2. | anne | *that (by hearer)* | akanne | *those (by hearer)* |
| 3. | arei | *that* | akekei | *those* |

with very similar temporal, locative, and discourse functions. Except for the (2) forms and some minor morphological processes we cannot go into here, the suffix -i indicates that no further delimitation of the noun phrase will follow. Absence of the -i requires the occurrence of a subsequent relative clause sentence. In this connection it is interesting that if one adds an -i to the complementiser bwa *that* the resulting form bwai is identical with the word for *thing, fact*.

We will see that the relative proforms also play an interesting role marking sentences embedded as noun phrase complements.

These forms also occur in a special type of equative used for the counterpart of some WH questions in English:

75. Te-raa ae ko-a tia-n noori-a?
ART-*what that you*-REALIS COMPLETIVE *see-it*

*What did you see?*

76. Antai are e-ka-uka te mataroa?
*who that he*-CAUSE-*open* ART *door*

*Who opened the door?*

77. Ningai ae e-a mate?
*when that he*-REALIS *die*

*When did he die?*

In these sentences the ae/are clauses are given information, the subject noun phrases. The question word is a predicate noun phrase. We might thus translate (76) more literally as *Is who, the one that opened the door?*

## VI. SOME PROBLEMS CONCERNING COMPLEMENTATION WITH BWA

As mentioned just above, KIR has a complementiser bwa, translated as *that* because it introduces embedded clauses. So, we can replace the object noun phrase te koaua *the truth* in (78):

78. I ata-a te koaua
*I-know*-TRANSITIVE ART *true*

*I know the truth.*

with a sentential noun phrase:

79. I-atai-a bwa ko-na roko ningabong.
*I-know-it that you*-IRREALIS *come tomorrow*

*I know you'll come tomorrow.*

But *that* is not always a good translation. For example the question (76) above can be embedded after bwa, as in (80):

80. E-a tia-ni kaoti-a bwa antai are e-ka-uka
*he*-REALIS COMPLETIVE *point-out-it that who that he-cause-open*

te mataroa
ART *door*

*He has pointed out who opened the door.*

But that could not be used for bwa here. Bwa has also been translated as because, just as Trukic pwe, *that*, often is too, along with the corresponding forms in other MC languages. We will consider this special role later in our discussion. There is one other apparently special role with the verb riki *become, come about, happen*:

81. E-na riki bwa te tia-moti
*he*-IRREALIS *become that* ART *judge*

*He'll become judge.*

The complementiser bwa appears to be followed by a simple noun phrase rather than an embedded sentence. It is possible to specify the subject more fully and the specified form occurs in extraposed position

82. E-na riki bwa te tia-moti Nakaa
*he*-IRREALIS *become that* ART *judge*

*Nakaa will be the judge.*

There is however an alternative analysis which can save our analysis of bwa as always an introducer of embedded sentences. We have seen that te tia-moti can be a predicate noun phrase. Under such an analysis Nakaa is the specified extraposed subject of the embedded sentence:

83. Te tia-moti Nakaa.

*Nakaa is the judge.*

So in (82) Nakaa may not have been extraposed from subject position in the higher clause. This will be clearer from the following example:

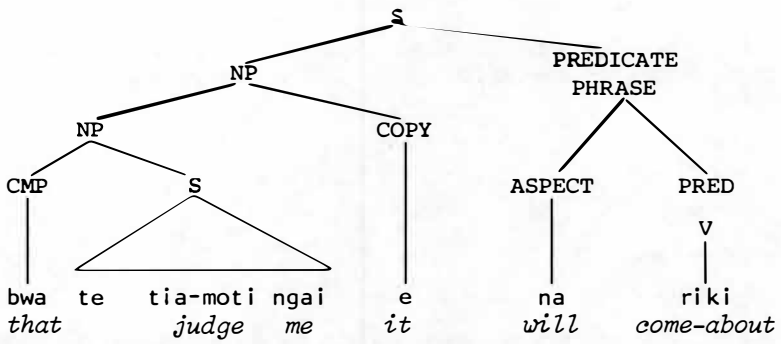
84. E-na riki bwa te tia-moti ngai
*me*

*I will become the judge.*

The first-person independent pronoun cannot have come from the higher clause because the subject marker left behind in the higher clause is third-person singular. The forms after bwa clearly constitute an embedded equative sentence. This also accounts for the 'new information' assertive status of te tia-moti.

But what then does the subject copy e- refer to in (84)? Our answer is that riki is more accurately translated as *come about* here and that e- is the predictable third-person copy. It is the trace of an extraposed third-person subject noun phrase, the entire clause bwa te tia-moti ngai. Prior to EXTRAPOSITION but subsequent to SUBJECT COPYING, the constituent structure for (84) would have been (85) below:

85.




*That I (am) judge will come about.*

The fact that there is a paraphrase of (84) with a first-person subject copy in the higher clause indicates that there may be a RAISING much like the Eng. one converting the unextraposed version of

86. It happened that I was nearby.

into

87. I happened to be nearby.

A precise formulation of RAISING awaits more detailed investigation. One type of RAISING, that involved here, is not raising in a literal sense. A copy of the lower subject replaces the original third-person subject copy of the higher clause. The bwa clause can remain intact with a possible condition that it not be a subject or object. Raising from lower subject to higher object position applies to structures underlying sentences like (88):

88. I-tangiri-a bwa ko-na nakomai
*I-want-it that you*-IRREALIS *come-here*

*I want you to come here.*

to generate

89. I-tangiri-ko bwa ko-na nakomai
*I-want-you that you*-IRREALIS *come-here*

*I want you to come here.*

We will return to RAISING in our discussion of the verbal auxiliaries in KIR.

It was noted above that bwa clauses, unlike that clauses in English, can remain intact after raising. If this kind of raising is merely a copying rule, such a phenomenon is hardly unexpected. But the condition that it not be a surface subject or object is surprising. Part of the evidence for this claim,

moreover, is shaky. Is it in fact true that extraposed subject noun phrases cease to be subjects? We cannot be sure at present since there is as yet no detailed study of subject and object properties in this language, nor of processes affecting them. What about object complements?

In KIR, many transitive verbs have a special object suffix when a singular object noun phrase is not specified. This suffix can refer to an animate being or to a proposition. Thus (90) has three renderings:

90. I-atai-a
*I-know-it/him/her*

*I know it/him/her.*

provided that the it refers to some previously mentioned proposition. If a singular object noun phrase is specified, then the ordinary transitive thematic suffix -a occurs:

91. I-ata-a te koaua
*I-know* ART *truth*

*I know the truth.*

92. I-ata-a te mwaane

*I know the man.*

But when a bwa clause appears in what should be object position, the verb has the unspecified object suffix:

93. I-atai-a bwa ko-na roko ningabong
*I-know-it that you*-IRREALIS *come tomorrow*

*I know you'll come tomorrow.*

The homophony of the two -a suffixes seems synchronically irrelevant, especially as they occur in different positions on the verb stem. One explanation of the fact that bwa complements are not treated as objects in surface structure is that, in fact, they are not surface objects. They may be like deep subject complements in that they are always extraposed. As in Eng., extraposition from object is far less obvious than extraposition from subject and this earlier caused some doubt as to its existence. The evidence for it in KIR seems stronger.

Bwa followed by an ordinary embedded sentence can never appear in sentence-initial position, e.g. as topic. Thus while complex topics like the following are not uncommon in traditional texts:

94. Ao te uea ae (e-) korakora ae Auriaria, e- karauii nano-ia
*and* ART *king that (he) great that he- comfort mind-their*
aomata
*people*

*Then the great king Auriaria comforted the hearts of the people.*

a bwa clause cannot occur sentence-initially:

95. *Bwa a-na maeka ma-ngaiira, e-a kakukureia-i
*that-they*-IRREALIS *live with-us it*-IRREALIS *pleased-me*

*That they would live with us pleased me.*

Various syntactic processes serve to prevent initial bwa.[2] The adverbial predicate noun phrase ibuki-n te-raa, *because of what, why,* precedes the extraposed bwa subject in the question:

96. Ibuki-n te-raa bwa e-a ngongo?
*because-of what that he*-REALIS *itch*

*Why is he itching?*

But the sentence following bwa can remain in initial position provided that it is de-subordinated by lowering the adverbial predicate:

97. E-a ngongo ibuki-n te-raa
*he*-REALIS *itch because of what*

*Why is he itching?*

Passive structures are also a means of avoiding initial bwa. Thus the passive (98) corresponds to an active sentence with bwa as its subject, but the bwa S is extraposed, as in (99):

98. I-ka-kukurei-aki bwa a-na maeka ma-ngaiira
*I*-CAUSE-*happy*-PASV *that they*-IRREALIS *live with-us*

*I was pleased that they were going to live with us.*

99. E-a ka-kukurei-a-i bwa a-na maeka ma-ngaiira
*it*-REALIS *cause-happy-me that they*-IRREALIS *live with-us*

*It made me happy that they lived with us.*

The bwa clause, no longer a subject, seems to be an oblique noun phrase bearing no major grammatical relation, much as adverbs. In fact, the clumsy (99) is likely to be phrased more simply as (100):

100. I-a kukurei bwa a-na maeka ma-ngaiira
*I*-REALIS *happy that they*-IRREALIS *live with-us*

*I'm happy that they're going to live with us.*

where the 'cause' relation is pragmatically inferrable from the combination of the emotional state and the prospect to which the state is a reaction. This kind of pragmatic relation has led grammarians of several MC languages to describe the bwa counterparts as 'because' conjunctions distinct from the complementiser. The same kind of reasoning would identify that in the Eng. version of (100) as a causal conjunction. Like that clauses in Eng. bwa clauses cannot be preceded by a preposition that would otherwise mark them clearly as oblique noun phrases.

However, bwa complements do occasionally have head noun phrases:

101. Ngkai ti-a bon ataa te koaua bwa bon akea
*now we*-REALIS ASSERT *know* ART *truth* ASSERT *not-exist*
Nei Nibarara iaon namwakaina.
*on moon*

*Now we are really sure of the fact that there is no Nei Nibarara on the moon.*

though they are frequently extraposed away from their heads:

102. E-a boni ka-nako-a naba ana taeka te nati-ni-uea
*she*-REALIS ASSERT *cause-go also her word* ART *daughter-of-king*

bwa e-na katietie-aki iroun te tei-ni-mmwaane arei
*that she*-IRREALIS *swing*-PASV *by* ART *boy that*

*The princess sent the message that she wished to be swung by that boy.*

One final phenomenon involving bwa is puzzling at first sight. Bwa is sometimes followed by are, which we identified earlier as a relative clause marker. Sometimes the are serves as a head for a referential entity while the bwa perhaps just marks embedding:

103. I-a noora bwa are ko-uoti-a
*I*-REALIS *see that that you-bring-it*

*I saw what you brought.*

which also involves the non-referential meaning, *I saw that you brought something*, hence, possibly, the bwa are co-occurrence. But this explanation will not work for (104) which has no referential entity to be marked with are:

104. Ti-kakoaua raoi bwa are a-bon tia-n roko
*we-believe firmly that that they*-ASSERT COMPLETIVE *come*

i-namwakaina i-Amerika
*at-moon Americans*

*We are convinved that the Americans have already arrived on the moon.*

The bwa are constructions occur in our data where either reference is intended or where a factive sense seems to reside. If the latter observation holds, the function of are as marking an assumed or 'given' relationship between a head noun phrase and its relative has made it eligible to mark presupposed truth, i.e. fact, for complement sentences.

## VI. EQUI DELETION IN KIR

Some verbs which take bwa sentential complements also occur with a different construction which appears to paraphrase the bwa clauses. For example, in the sentence

105. A-bon aki tangiri-a bwa a-na kautia te unnaine
*they*-ASSERT *not want-it that they*-IRREALIS *awake* ART *old-woman*

*They didn't want to wake up the old woman.*

the subjects of the higher and lower clauses have the same referent. The verb tangira does not require that this be so, but, when such coreference occurs, an alternative structure is available as in (106):

106. A-bon aki tangiri-a ni kautia te unnaine
*they*-ASSERT *not want-it awake* ART *old-woman*

*They didn't want to wake up the old woman.*

Sentence (106) has ni instead of bwa and what follows ni is not a clause with subject copy, aspect, and the rest of the predicate phrase but simply the predicate phrase starting with the verb. Nevertheless, KIR speakers know that

the persons referred to in the higher clause are also the persons involved in the possible waking of the old woman.

We therefore assume that EQUI NOUN PHRASE DELETION can apply to structures like that underlying (105), deleting the coreferential lower subject noun phrase. Such reduced clauses in Eng. may be marked for aspect and voice, though not tense. In KIR, aspectual marking cannot occur in such constructions but passive voice marking on the verb (with -aki) is very common. The EQUI process can apply to intransitive verbs like nako *go*, so that

107. E-a nako Ribwa bwa e-na kooni taian reemon
*she*-REALIS *go* *that she*-IRREALIS *squeeze* *some lemons*

*Ribwa went to squeeze some lemons.*

can become:

108. E-a nako Ribwa ni kooni taian reemon.
*she*-REALIS *go* *to squeeze* *some lemons*

with the same basic meaning.

As in Eng., for verbs like imwanonoa, *force*, which take three noun phrase arguments, the controller for the deletion is not the higher subject but an animate noun phrase object of the higher verb. The shared semantic characteristic of all such EQUI-triggering verbs is that they are all future-oriented — the event represented in their lower S is always future with respect to the time reference of the higher verb. The function of EQUI is to reduce the hierarchical structure by purging the lower clause of constituents whose sense is completely recoverable from the higher clause. This is most obviously the case for verbs which allow only the n/ni[3] construction, i.e. for which EQUI is obligatory. This kind of verb requires the deep subject of its lower sentence to be coreferential with its own subject. This is in contrast to verbs like tangira *want*, noora *see*, nako *go* for which such coreference is optional, and for which EQUI need not apply even if the structural description is met. Verbs like kataia *try*, kona *be able, can*, moana *begin*, tataneia *used to*, which occur only with n(i) seem more like modal auxiliaries, and, for some of them, some process other than EQUI may be involved. We will discuss some of these cases shortly.

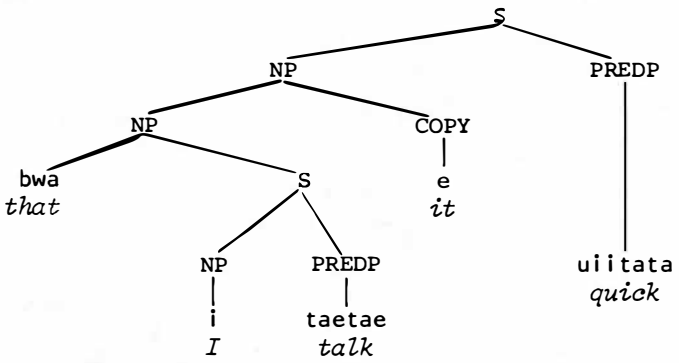
Just as interesting from a cross-linguistic point of view is the fact that some verbs we might expect to allow, if not require, EQUI may perhaps occur only with full bwa complements. These include tuanga *order, tell* and kataua and kariaia *allow*. However this, perhaps understandably, appears to vary among speakers and it requires further study.

According to Cowell, the particle n(i) is really four particles — an infinitive marker also used to form adverbial phrases, a genitive proposition (te umwa n roronga *the house OF young men*), an agent preposition with inanimate agents (e-karaoaki n te atibu *it was done by the rock*), and a locative preposition (n te umwa, *in the house*). We might say instead that the particle appears to be a general linking element and that the interpretations Cowell assigns arise from their syntactic and pragmatic contexts. The particle allows additional non-nuclear noun phrases to be added to clauses and identifies them as 'circumstances'. Likewise it allows the collapsing of an object clause with the higher clause and, especially where the other clause is not a term (subject or object), it reduces the clause to a n VERB structure that can be translated with an English adverb.

## VII. A RAISED CONSTRUCTION WITH N(I)

We now examine one such structure in more detail. Assume an original Topic-Comment underlying structure in which the Topic is ngai, *I*, and the major assertion of the Comment is that my speech is rapid. After whatever processes copy the Topic into subject position and then delete it after SUBJECT COPYING, and after bwa has been introduced before the clause complement, we would have a structure corresponding more or less to (109) below:

109.



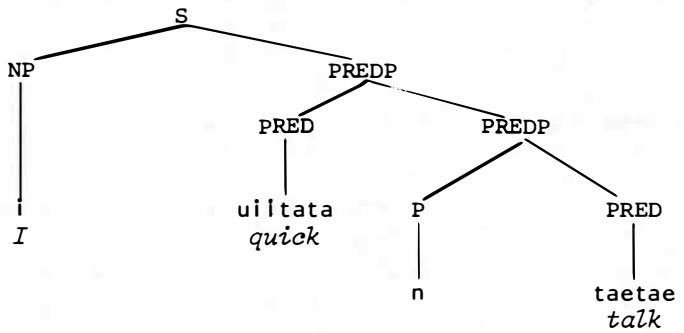

EXTRAPOSITION can now be applied to shift the bwa clause to the position after the predicate uiitata:

110. E-uiitata bwa i-taetae
*it-quick that I-talk*

*I speak quickly*

Alternatively, before the subject copies have been attached to the predicates,[4] we can instead RAISE the lower subject to higher subject position and shift the predicate phrase, minus any aspect marking, into a n(i) phrase at the end:

111.




I-uiitata n taetae
*I-quick talk*

*I speak quickly*

It is also possible to lower the higher predicate into a n(i) adjunct:

112. I-taetae n uiitata.

*I talk quickly.*

The n(i) phrase in (112) is the form sometimes described as adverbial.

## VIII. SOME AUXILIARY FORMS IN KIR

We can use the following to represent the order and possible constituents of the pre-verb part of a predicate phrase in a predicative sentence. The formula is not quite correct because the positions of the assertive particle, aspectual marker, and negative are somewhat variable and, in the case of auxiliaries, several may co-occur, each followed by a n(i).

113. COPY-(ASSERT) (ASPECT) (NEG) (AUX n(i))

We are concerned here witn the part we have marked as AUX n(i). KIR has a fair number of these auxiliary forms, all of which seem to express primarily modal or aspectual notions. They include tia, which we have labelled 'completive', bwaane, *be all*, riai *can, trust*, tabe and kume, which are used as a kind of progressive/durative markers, kataia, *try*, kan, *want*, kaan *almost*, toki *stop*, taatanei *used to*, and others.

We want to determine more precisely the nature of the relation between the auxiliary forms and the verbs following. Why do just these forms and none of the others — the copy, assertive and aspect markers, or the negative — require a n(i) to follow them? We have seen that combinations of verbs formed as a result of EQUI are separated by n(i). This is certainly likely to be what happens with kataia *try*, which looks like a verb anyway, with its ka-...-a causative marking.

But what of the 'completive' tia, riai *can, must*, and the 'progressive/durative tabe and kume? If they are verbs, they have no objects in surface structure. Moreover there are some interesting paraphrase relations to account for. Why should (114) and (115) below have almost exactly the same meaning?

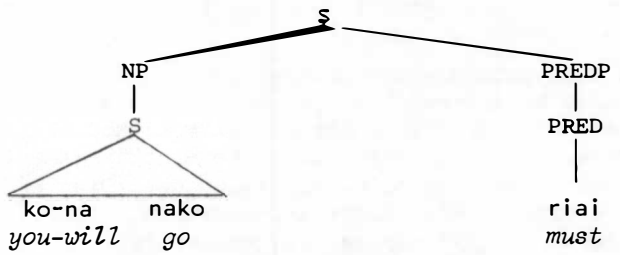
114. E-riai bwa ko-na nako
*it-must that you*-IRREALIS *go*

115. Ko-riai n nako
*you-must to go*

*You must go.*

Sentence (114) contains a riai in an environment quite distinct from any suggested by the formula of (113). The form riai has no n(i) following it; nor is a main verb the next major contentive. The bwa sentence complement has obviously been EXTRAPOSED. If this is so, then riai would appear to be a verb which has a sentence complement as subject. A reasonable underlying structure might be (116):

116.




Structurally this is just the kind of construction we posited for the uiitata *(quick)* construction earlier. SUBJECT COPYING can apply, followed by EXTRAPOSITION, yielding (114). The bwa complementiser has to be inserted since the clause has retained its integrity as a clause. On the other hand, we could

instead RAISE the subject, ko *you*, of the subject complement sentence into subject position in the main clause and attach the predicate phrase nako (without its aspectual markers) to the higher predicate phrase by means of n(i). So riai appears underlyingly to be an intransitive verb. We can check this with another criterion — is it possible to transitivise riai *must, can* with causative affixation (ka-...-a for the third person singular)? The form kariaia is indeed a verb and it means *allow*, a fairly plausible causative for a verb having as one meaning *can, be able*.

Not all of the forms allow the possibility of EXTRAPOSITION. As with the English auxiliaries may, can, should, they do not occur in sentence-final position either. Gundel and Jacobs have claimed this as a general characteristic of verbs whose primary role is to express modality, since the major assertion is not in such a verb but in the embedded clause which is its subject noun phrase; many languages position the most assertive constituents, the newest information, in sentence-final position. For tia, tabe, and kume, only the RAISING alternative is available. In fact, these three forms also have causatives corresponding to them: katiaa *cause to be completed*, katabea *cause to be busy*, and kakaumea, *bother*, all of which have at least diachronic relation to the auxiliary verbs. The intransitive verb tia *finished, complete*, is very common:

117. E-a tia te umwa
*it*-REALIS *finish* ART *house*
*The house is finished.*

Even if we ignore the connections with causative verbs and intransitives, it seems economical, in terms of both processes and categories in KIR, to regard these auxiliary forms as main verbs. We need no special processes to generate such forms, except perhaps for a crosslinguistic restriction on the position of non-assertive predicates. And the semantics of these forms seems more accurately represented in such an analysis.

## IX. A FEW CONCLUSIONS

KIR is a very 'verby' language, using verbs where other languages might have more numerous and specialised categories. It has very few 'genuine' prepositions, the others being verbs or nominals. Forms translated into Eng. as quantifiers — *all, many* etc. are verbs, as are the KIR equivalents of many Eng. adverbs. As in other Pacific languages, many noun forms can, without modification, be used as ordinary verbs.

A language with comparatively few categories and many verbs would seem to be a very hierarchically structured one with many clauses. Presumably such a language might present problems for memory capacity although its semantics would be more transparent. Gorbet has pointed to a Hokan (Amerindian) language, Diegueño, as an extreme example of a 'verby' language. Diegueño makes use of an elaborate system of case-marking and switch-reference suffixation which forestalls such problems. In KIR, syntactic complexity is reduced by a few significant reduction and restructuring processes — EQUI, RAISING, and EXTRAPOSITION — along with an industrious exploitation of proforms, assertive/referential contrasts, as in relative and equatives, and by a very productive morphology. We suspect that the other nuclear MC languages are similar in this respect.

## NOTES

1. It seems likely that such Topic-Comment structuring should be represented in underlying semantic structure along with some kind of marking of the Given-New informational status of the various semantic constituents. Furthermore it may in fact be preferable to posit underlying predicate-initial structuring and thereby avoid having to posit obligatory extraposition for all subjects. Object-marking on predicates is far less regular for plural animate objects and it is possible that the phenomena noted here are in fact synchronic relics of a significant diachronic shift. We are currently exploring these promising possiblities.

2. We have encountered a few cases of sentence-initial bwa but all of them involved a following noun construction buki-n *base of* (literally), the meaning being *because, for the sake of.*

3. The form of the n(i) is phonologically conditioned by the following segment. The particle is ni before labial and velar consonants, including nasals, and before i. Otherwise the form is n.

4. The precise formulation of these processes still requires much investigation. This kind of RAISING appears to require that the raised noun phrase have the same referent as the Topic. It is fairly clear that the SUBJECT COPYING rule will have to apply after RAISING although we have assumed the reverse to keep our exposition fairly simple. Furthermore it may be possible, even desirable, to reformulate this RAISING as a LOWERING rule.

## BIBLIOGRAPHY

BINGHAM, Hiram

1922 *Outlines of a grammar of the Gilbert Islands language.* Rongorongo, Beru, Gilbert Islands: London Missionary Society. (Originally compiled in 1861 in MS form only.)

CHOMSKY, Noam

1965 *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

COWELL, Reid

1951 *The structure of Gilbertese.* Beru, Gilbert Islands: Rongorongo Press.

DOWNING, Bruce T.

n.d. English relativization. Mimeo. Department of Linguistics, University of Minnesota.

GORBET, Larry

1974 Relativization and complementation in Diegueño. Doctoral dissertation, University of California at San Diego.

GUNDEL, Jeanette K.

1974 The role of topic and comment in linguistic theory. Doctoral dissertation, University of Texas at Austin. Reproduced by Indiana Linguistics Club.

GUNDEL, Jeanette K. and Roderick A. JACOBS

1980 Why "seem" and "be" aren't what they seem to be. *Centrepoint* 4/2: 115-123.

JACOBS, Roderick A.

1973 Syntactic compression and semantic change. In Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser, eds *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, April 13-15, 1973 1973,* 232-241. Chicago Linguistic Society.

1975 Promotion and thematisation processes in English, or, How to get a head. In Robin E. Grossman, L. James San, and Timothy J. Vance, eds *Papers from the Parasession on functionalism, April 17, 1975,* 223-231. Chicago Linguistic Society.

1975 *Syntactic change: a Cupan (Uto-Aztecan) case study*. Berkeley: University of California Press.

1976 A passive continuum in Austronesian. In Sandford B. Steever, Carol A. Walker, and Salikoko S. Mufwene, eds *Papers from the Parasession on diachronic syntax, April 22, 1976,* 118-125. Chicago Linguistic Society.

1976 Syntactic reconstruction and the comparative method: a Uto-Aztecan case study. In Mohammad Ali Jazayery, Edgar C. Polomé, and Werner Winter, eds *Linguistic and literary studies in honor of Archibald A. Hill,* vol.1: *General and theoretical linguistics,* 165-174. Lisse: Peter De Ridder Press.

1977 Syntactic categories and predictions in Gilbertese. Paper presented to the Linguistic Society of America Summer Meeting, Honolulu, Hawaii.

1980 Towards a semantic grammar of English. MS.

1983 Relatives, equatives, and information structuring. See pp. 37-51 in this volume.

SOHN, Ho-min

1973 Relative clause formation in Micronesian languages. *Oceanic Linguistics* 12:353-392.

STAHLKE, Herbert F.W.

1976 Which that. *Language* 52/3:584-610.

SUGITA, Hiroshi

n.d. Lagoon Trukese grammar. MS. University of Hawaii.

# ON THE VERBAL NATURE OF KIRIBATI bane 'ALL'

Sheldon P. Harrison

## 1. INTRODUCTION[1]

In Kiribati (Gilbertese) (KIR) the universal quantifier is encoded by the verb bane *exhausted, used up, all* in what can be termed a clause chain construction. The morphology and syntax of bane, when used as a quantifier, varies, however, depending on the surface status of the nominal in its scope in the clause encoding the propositional content of the sentence (that is, whether that nominal is surface subject, object, or oblique). Ignoring some minor patterns, if the quantified nominal is surface subject, bane appears as an intransitive verb initial in its chain; if non-subject, bane appears as a superficially 'causative'[2] verb non-initial in its chain.

The behaviour of bane as a quantifier follows from the interaction of constraints on coreference in clause chains and from the functions of causative derivation as a transitivising mechanism. It will be argued that while non-initial clauses with causative bane resemble other clause chain constructions over which a coreferential subject constraint holds, these clauses are in fact subjectless.

## 2. SURFACE GRAMMATICAL RELATIONS AND CLAUSE-CHAINING IN KIR

The surface grammatical relations subject and object are encoded in KIR verbal clauses by means of proclitic subject markers (referencing the person and number of the subject) and verbal object suffixes (referencing the person, number, and, in the third person plural, animacy/humanness of the object)[3]:

1. I a kua.
   SM ASP *tired*
   *I am tired.*

2. A matuu ngkai ataei.
   SM *sleep now children*
   *The children are sleeping now.*

3. I nooria ngkoaananoa.
   SM *see*-OM *yesterday*
   *I saw him yesterday.*

---

Byron W. Bender, ed. *Studies in Micronesian Linguistics*, 491-501. *Pacific Linguistics*, C-80. 1984.

4. E kairiia ataei nakon te titoa.
   SM *take*-OM *children to* ART *store*
   *He took the children to the store.*

5. A aki tangirai.
   SM NEG *like*-OM
   *They don't like me.*

6. E karea te kamea te ataei n te atiibu.
   Sm *throw*-OM art *dog* ART *child with* ART *stone*
   *The child threw a stone at the dog.*

Subject markers are obligatory in all verbal clauses except imperative clauses and non-initial clauses in a chain. In the analysis of KIR assumed here, object marking is a defining characteristic of the category transitive verb.

In their surface syntax what are here termed clause chains bear some resemblance to English sentences with infinitival adjuncts. They consist of a series of clauses linked by the morpheme ni (n before vowels and apical consonants):

7. Ko a inaaomata n oki.
   SM ASP *free* ni *return*
   *You are free to return.*

8. E tei n iango.
   SM *stand* ni *think*
   *He stood thinking.*

9. E a kabwakaa waana n nako n akawa.
   SM ASP CAUS-*fall*-OM *canoe-his* ni *go* ni *fish*
   *He took down his canoe to go fishing.*

10. E rawea atuuna n taua ma rarikin waana.
    Sm *catch*-OM *head-his* ni *hold*-OM PREP *side-of canoe-his*
    *He caught its head and held it against his canoe.*

11. A bon tia n roko I-Waruwaru ni maeka ibuakoia.
    SM ADV *finish* ni *come W.* ni *dwell among-them*
    *The Waruwaruans came to dwell among them.*

12. E kaea ni kanakoa ni batiboia n tiringnga ni kamatea.
    SM *chase*-OM ni *take*-OM ni *beat*-OM ni *slay*-OM ni *kill*-OM
    *He chased him and got him and beat him and slew him and killed him.*

Clause chains contain a single subject marker and modal/aspectual proclitic, drawn from an optional and mutually exclusive set including a 'immediate relevance', na 'irrealis', and nang 'immediate future', appearing before the first verb in the chain. In one sense, then, the entire chain has a single surface subject, typically the nominal that would appear as surface subject of all 'links' in the chain were it decomposed (for example, the same agent for all five chained clauses in 12.).

Some apparently chained constructions show coreferentiality between the subject of the ni-clause and the object of the preceding clause:

13. I namwakina te raanti ni buti.
SM *feel*-OM ART *launch* ni *move*

*I felt the launch move/moving.*

Under the view of KIR syntax assumed here such examples are not clause chains, however. Rather, they represent another construction type in which, in this instance, the phrase te raanti ni buti *the launch moving* is a noun phrase. The same 'participial ni' construction is found in sentences like:

14. Maanra ngkoe ni mwmwakuri?
*duration-what you* ni *work*

*How long did you work?*

Some ni-clause occurrences of bane *all* are open to a 'participial ni' analysis (see Section 6).

Some of the most frequently occurring chain-initial predicates are operators (over clauses) whose semantics preclude actant subjects:

15. I rangi ni kukurei n anne.
SM *very* ni *happy* ni *that*

*I'm very happy with that.*

16. I a tia ni mwmwakuri.
SM ASP *finished* ni *work*

*I've finished working.*

Rangi *very* is found only in clause chains. While tia *finished* does occur in isolation, its subject is always an event, never an individual:

17. E a tia te mwmwakuri.
SM ASP *finished* ART *work*

*The job has been finished.*

While the semantics of KIR clause chaining have yet to be investigated in detail, enough information is available to enable some conclusions to be drawn regarding the behaviour of bane as a quantifier in a clause chain construction.

## 3. SYNTAX AND SEMANTICS OF bane

The verb bane, in isolation out of a clause chain, is a stative whose surface subject is a semantic patient:

18. E a bane aia tianaki.
SM ASP *exhausted their provisions*

*Their provisions ran out.*

19. E aki baabane ana iango.
SM NEG RED-*exhausted his thought*

*He wasn't running out of ideas.*

(lit. *His ideas were not exhausted.*)

As with all other KIR statives, the valency of bane can be increased by adding the causative prefix ka- and appropriate object marking to derive a controlled transitive verb kabanea *to use up, run out of something*, with a typically agent subject:

20. E a kabanea korakorana.
SM ASP CAUS-*exhausted*-OM *strength-his*

*He had used up his strength.*

21. E bon aibabaia naba ni kabanea aitoana.
SM ADV *pull*-OM ADV ni CAUS-*exhausted*-OM *rival-his*

*He pulled at it to exhaust his rival.*

The most frequent use of bane, however, is to lexicalise the universal quantifier. In this use, it occurs in a (minimally) two member clause chain, as in:

22. A bane ni matuu iruwa.
SM *all* ni *sleep stranger*

*All the strangers were sleeping.*

23. A bane ni mamaeka ibuakonikai.
SM *all* ni RED-*dwell in-midst-of-tree*

*They all lived in the forest.*

24. Kamw na bane ni uotii kanamii.
SM MOD *all* ni *bring*-OM *food-your*

*You will all bring food.*

25. A bane ataeinimwaane ni kani kabuta te kaa.
SM *all* *boy* ni *want drive* ART *car.*

*The boys all want to drive the car.*

Clause chains in which bane functions as a universal quantifier are of four types, summarised in the following table:

Table 1: Quantifier bane constructions

| Type | Initial clause | ni-clause | Scope of ∀ | Examples |
|---|---|---|---|---|
| A | bane | proposition | surface subject | 22-25 |
| B | proposition | ka-bane | non-subject | 26-28 |
| C | ka-bane | proposition | surface object | |
| D | proposition | bane | surface oblique locative | |

Type A (henceforth SUBJECT-SCOPE bane CHAINS) and type B (henceforth NON-SUBJECT SCOPE bane CHAINS) are the most frequent bane constructions. The scope domains of types C and D are subdomains of Type B. Type C, in other than elicited examples, is attested only in the KIR Bible. Type D is more common, but can always be replaced by the corresponding type B construction. Types C and D will be treated in more detail in Section 5. and 6., respectively.

The salient differences between subject and nonsubject scope bane chains (these latter illustrated in 26. - 28. below) are then:

i) the relative position of bane in the chain
ii) the causative morphology (the prefix ka-) of non-subject scope bane
iii) the surface status of the nominal in the scope of bane

26. E nooriia raona ni kabane.
SM *see*-OM *friend-his* ni CAUS-*all*

*He saw all his friends.* (object scope)

27. Te uea e karaoa te amwarake ae e abwaabwaki nakoiia aomata
ART *king* SM *make*-OM ART *meal* REL SM *big* *to-them person*

ni kabaneiia.
ni CAUS-*all*-OM

*The king made a great feast for all the people.* (benefactive scope)

28. E rineaki irouiia kaain te kaawa ni kabane.
SM *choose*-PASS *by-them people-of* ART *village* ni CAUS-*all*

*He was chosen by all the people of the village.* (passive agent scope)

An account of the syntax of subject scope bane chains can be given in terms of:

i) constraint on the grammaticalisation of operators
ii) constraints on the recoverability of subjects in clause chains

All operators in KIR that are encoded as lexical verbs normally surface in a clause chain in a position preceding the grammaticalisation of the propositional content, whether these operators be modal, aspectual, or argument operators like the universal quantifier:

29. A riai n nako Bairiki n te bong aei.
SM *must* ni *go* B. PREP ART *day this*

*They must go to Bairiki today.*

30. A a tia n nako Bairiki n te bong aei.
SM ASP *finished* ni *go* B. PREP ART *day this*

*They've gone to Bairiki today.*

31. A bane n nako Bairiki n te bong aei.
SM *all* ni *go* B. PREP ART *day this*

*They all went to Bairiki today.*

Further, the person and number of the surface subject of the clause encoding the propositional content must always be referenced in a subject marker appearing initial in the clause chain. The subject must then be coreferential for all propositions in the chain, so that it need only be lexicalised once for the entire chain. In the view adopted here, with one exception, no propositional subject recoverability problem arises when the first clause of the chain is a modal or aspectual operator since these operators have no logical arguments and, hence, can have no surface subjects distinct from that of a following propositional clause. The subject marker then uniquely references the surface subject of the clause encoding the proposition, no matter what operators intervene on the surface, although grammatically it remains the subject of all clauses (operator or propositional) in the chain.

Recoverability problems arise only in the case of quantifiers grammaticalised as verbs in a clause chain, since the element on which the quantifier operates (the nominal in its scope) will typically be an argument of the proposition encoded in the chain. The subject marker must then fill a dual role, referencing the subject of the propositional clause and the nominal in the

scope of the quantifier. Only when these nominals are identical is the grammaticalisation of the universal quantifier straightforward, resulting in type A subject-scope bane constructions.

In a sentence like:

32. Ti bane ni kang iika.
SM *all* ni *eat*-OM *fish*
*We all ate the fish.*

logically either argument of *eat* (*we* or *fish*) might be in the scope of the universal quantifier, but only the interpretation with subject (*we*) scope is possible since it is only in this interpretation that the subject marker uniquely determines both the scope of the quantifier and the subject of the propositional clause. Were an interpretation of 32. with object scope required, the following sentence would result:

33. Ti kang iika ni kabane.
SM *eat*-OM *fish* ni CAUS-*all*

*We ate all the fish.*

where bane no longer occupies normal (chain-initial) operator position and, moreover, has acquired causative morphology.

To make clear the relationship between subject and non-subject scope constructions like 32. and 33., it may be opportune to view the latter as the result of a process that might be termed subject demotion (without wanting to claim that these constructions are necessarily derived by some rule or process of subject demotion in some particular theoretical framework), which has the effect of grammaticalising a propositional clause non-subject in the scope of bane as surface object of the bane clause. Since only overt transitive verbs have surface objects, bane then acquires the causative morphology appropriate to transitivised statives in KIR. (Arguments for the object status of the nominal in the scope of surface kabane are given in Section 4.) The process is considered to be subject demotion in view of the fact that the nominal in the scope of bane, otherwise referenced by the subject marker of the clause chain of which the surface bane clause is the first member, can be viewed as the surface subject of the bane clause, as in 22. - 25.

## 4. SURFACE GRAMMATICAL RELATIONS IN ni kabane CLAUSES

Grammaticalising propositional non-subjects in the scope of bane as surface objects of kabane transitive clauses obviates the subject recoverability problem that would arise in clause chains involving universally quantified non-subject nominals (where the subject of bane, that is, the nominal in its scope, is not also the subject of the propositional clause). No subject recoverability problem arises in such clause chains (for example, 26. - 28.) since the quantified nominal is surface object, not surface subject, of the kabane clause.

The morphological causativity of kabane in such cases is, however, insufficient to demonstrate the syntactic transitivitiy of kabane clauses, given that KIR, like other MC languages, allows intransitive causatives, as in:

34. E kamaamate te iti.
SM CAUS-RED-*die* ART *electricity*

*Electricity is dangerous.*

35. A tabe ni kakarabwakau naakai.
SM *busy* ni RED-CAUS-*knowledgeable these-men*

*These men are conversing.*

where the single overt nominal argument of the causative clause, te iti *electricity* and naakai *these men*, respectively, is surface subject of the clause (chain) and where no surface object is present.

The syntactic transitivity of KIR clauses, and thus the presence of an object, can usually be demonstrated by the potential for passivisation, as in:

36. I kaamwarakeiia ataei.
SM CAUS-*eat*-OM *children*

*I fed the children.*

37. A kaamwarakeaki ataei irouu.
SM CAUS-*eat*-PASS *children by-me*

*The children were fed by me.*

Passives of kabane quantifier clauses are disallowed, however, since such structures would automatically violate the coreferential subject constraint on clause chains. That is, the passive subject of the kabane clause would remain non-subject of the propositional clause, thus re-establishing the coreferentially problem that motivates subject demotion (Type B bane constructions). (Note that passivisation of a transitive propositional clause whose object is universally quantified yields a subject-scope bane chain, as in:

38. A bane ni kanaki iika iroura.
SM *all* ni *eat*-PASS *fish by-us*

*All the fish were eaten by us.*

a possible, but rarely encountered alternative to 33. above.)

The most obvious test for object status of a nominal in KIR is the presence of an object marker on the verb. Quantifier kabane is somewhat idiosyncratic with respect to object marking, however. In some cases object markers do appear on kabane (as in 27. above), but in most cases they do not. It would appear, with the exception of Type C bane constructions, that kabane references only animate surface objects, and then only optionally:

39. Ti kinaiia aomata akekei ni kabane(iia).
SM *know*-OM *person those* ni CAUS-*all*-(OM)

*We know all those people.*

40. Ti bon atai baikanne ni kabane(*i).
SM ADV *know*-OM *those-things* ni CAUS-*all*-(*OM)

*We know all those things.*

The sporadic nature of object marking with kabane suggests that a reanalysis of these constructions may be under way, a hypothesis that will not be explored here. The fact that such object marking occurs at all does, however, suggest that the nominal referenced by the object marker (the nominal in the scope of bane) is surface object of a transitive kabane clause.

Somewhat more vexing is the question of the status of the surface subject relation in non-subject scope kabane clauses, where the nominal that would be expected to be grammaticalised as subject (the nominal being quantified) is in fact grammaticalised as object. To the extent that the nominal referenced by

the subject marker appearing initial in a clause chain is subject of the entire chain, then ni kabane clauses do have subjects. It should be pointed out, however, that ni kabane clauses do not have subjects in the same sense as do 'typical' non-initial clauses in a clause chain, whose subjects are grammaticalisations of a particular argument of the proposition realised as the non-initial chained clause in question. bane (or kabane) clauses are semantically one-argument clauses, that argument being the nominal in the scope of the quantifier. The superficial transitivity of kabane clauses is a purely syntactic response to the problem of specifying the scope of the universal quantifier. As already pointed out, the nominal in the scope of the quantifier is grammaticalised as subject of bane unless this grammaticalisation results in a subject coreferentiality conflict within the chain, in which case it is grammaticalised as object of a superficially causative and transitive kabane. The 'causative' prefix ka- with bane can thus be viewed as serving at least two distinct functions:

i) the lexical (derivational) function of increasing the valency of stative bane *exhausted, run out* to permit the expression of a controller (agent)

ii) the syntactic function of permitting a nominal otherwise grammaticalised as subject of quantifier bane *all* to be grammaticalised as non-subject (object) in cases where the coreferential subject constraint on clause chains would otherwise be violated

The fact that kabane, though syntactically transitive, remains semantically intransitive can be used to account in part for the fact that kabane vacated the typical operator position, initial in the clause chain. If one considers this position to be restricted to predicates that are either without arguments (as in the case of modal/aspectual operators) or to predicates for which the subject of the clause chain is an argument (as in the case of non-operators of subject-scope bane), then non-subject-scope kabane will not occupy the typical operator position since the subject of the chain in which it occurs is not the nominal in the scope of bane. There is, however, one set of circumstances in which kabane does occur in operator position in the chain, the Type C bane construction. The properties of this construction are considered briefly below.

## 5. CHAIN-INITIAL kabane

As noted earlier, chain-initial kabane is a rare construction attested, in data available to me, only in the KIR Bible:

41. Kamw na kabaneiia ateainimwaane akana a bungiaki n teweiia
SM MOD CAUS-*all*-OM *boys* REL SM *bear*-PASS ni *throw*-OM
nako nanon te karanga.
*to* *inside-of* ART *river*
*You shall throw all the boys who are born into the river.*

42. Ao Mote e kabanea te bootaki ake tibun Iteraera n ikotiia.
*and M.* SM CAUS-*all*-OM ART *group* REL *descendant-of I.* ni *gather*-OM
*And Moses assembled all the congregation of the children of Israel.*

43. E na kabane am kai ma uaan abam te rokati ni bwain.
SM MOD CAUS-*all your tree and fruit-of land-your* ART *locust* ni *own*-OM
*The locust will possess all your trees and the fruit of your land.*

Object marking is much more frequent, if not obligatory, with chain-initial kabane than with ni kabane (43. being unique in my data in not exhibiting agreement). Moreover, chain-initial kabane constructions are fully acceptable only when the quantified nominal is surface object of the propositional clause. Thus 44., with a passive agent in the scope of bane, and 45., with a human goal in the scope of bane, were rejected by informants. 46., identical to 45. except for the surface position of unimwaane *elder*, was considered marginally better but still odd:

44. *E kabaneiia ni kabutaki te kaa irouia.
SM CAUS-*all*-OM ni *drive*-PASS ART *car by-them*
*The car was driven by them all.*

45. *I kabaneii ni kaota te reta aarei nakoiia unimwaane.
SM CAUS-*all*-OM ni *show*-OM ART *letter that to-them elder*
*I showed that letter to all the elders.*

46. ?*I kabaneiia unimwaane ni kaota te reta aarei nakoiia.
SM CAUS-*all*-OM *elder* ni *show*-OM ART *letter that to-them*
*I showed that letter to all the elders.*

The only obvious distinguishing characteristic of chain-initial kabane is its restriction to object scope. Under this restricted set of conditions, one might hypothesise, a tighter form of chaining results in which, superficially, both subject and object of each link in the chain are shared by the entire chain. This would not be the case if the scope of bane were the subject of the propositional clause (as no kabane clause would result), nor if it were an oblique nominal in a transitive propositional clause (since that clause and the kabane clause would have distinct objects), nor if it were an oblique nominal in an intransitive clause (since only the kabane clause would have an object). While this construction remains anomalous in any case, its occurrence does appear to be a function of the complete argument sharing noted above.

## 6. ni bane CLAUSES

Similarly anomalous are Type D constructions, with bane that is not chain-initial, as in:

47. E na tauaki te boti ibukin te roro rimwii ni mwaneaba ni bane.
SM MOD *hold*-PASS ART *place for* ART *generation after* PREP *meeting-house*
ni *all*
*A place will be reserved for future generations in all meeting houses.*

48. A ukeria iaon Tarawa ma aban Tungaru ni bane.
SM *search*-OM *on* T. *and land-of T.* ni *all*
*They searched for him on Tarawa and all the lands of Tungaru.*

49. E kaueaaki iaon Tarawa ni bane.
SM CAUS-*king*-PASS *on T.* ni *all*

*He was made king on all of Tarawa.*

in which:

i) bane is not in the typical operator position

ii) the subject of bane (the nominal in its scope) is not the surface subject of the propositional clause

It is significant that this construction is restricted, in my corpus, to locative complements. As has frequently been pointed out, such complements are 'peripheral' in more than one sense of that term. Most significantly, their occurrence is not a function, in most cases, of the inherent semantics of the predicate of the clause in which they occur. I would suggest, tentatively, that a verbal operator on a peripheral nominal of this sort does not form part of a clause chain with the proposition to which that nominal is peripheral, and is therefore not subject to the usual constraints on coreference governing KIR clause chains. Type D bane constructions, then, resemble the 'participial ni constructions' considered in Section 2.

Some evidence that the juncture between locative nominals and the propositions of which they might be taken to be arguments is in fact weak is provided by constructions like the following:

50. N te itera meaang ao a rineaki taan tautaeka mai buakoia
PREP ART *half north and* SM *choose*-PASS *doers rule from among-them*

utuun inaaomata.
*family-of free*

*In the northern half, officials were chosen from among the chiefly families.*

in which a locative phrase is left co-ordinated with the rest of the sentence, in the same way as distinct clauses, with distinct subjects, would be co-ordinated.

## 7. CONCLUSION

Though KIR bane *'all'* is semantically an operator on nominals, it is governed by the same syntactic constraints as operators on propositions; crucially, it forms part of a clause chain with the proposition of which the nominal in its scope is an argument. Clause chaining in KIR is governed by a like-subject constraint on all surface verbs in the chain. Where this constraint would be violated for bane (when its subject, the nominal in its scope, is not surface subject of the propositional clause) bane appears as a transitive causative. Only when the nominal in the scope of bane is peripheral (a locative) can the like-subject constraint be violated. In such cases it can be argued that the peripheral nominal in fact stands outside the propositional core. Operators on peripheral nominals thus will not form part of a chain but will be linked to their operand in a 'participial-ni' construction.

## NOTES

1. An earlier version of this paper was read at the annual meeting of the Australian Linguistic Society, Monash University, Melbourne, August 1980. Thanks to Bill Foley, Alan Dench, and Monty Wilkinson for comments. The usual disclaimers apply.

2. The term 'causative' is usually applied to verbs with the prefix ka- and will be used in this sense here, although the semantic suggested by that term is appropriate only to some functions of KIR ka-.

3. Most KIR transitive verbs distinguish third person singular pronominal object and expressed nominal object forms:

   E noora te aine. — *He saw the woman.*

   E nooria. — *He saw her.*

   No such contrast is found in the third person plural however. There, a somewhat complex contrast involving animacy/humanness holds, the details of which are irrelevant here:

   E nooriia aomata akekei. — *He saw those people.*

   E noorii waa akekei. — *He saw those canoes.*

   A small class of 'cognate object' verbs, including perception verbs like ongo *to hear* and psych verbs like mwaninga *to forget* have no object marking. These verbs are considered, under the analysis assumed here, to be superficially intransitive and are, in any case, irrelevant here.

# BIBLIOGRAPHY

ABBREVIATIONS

CLS = Chicago Linguistic Society
*LG* = *Language*
*OL* = *Oceanic Linguistics*
*PL* = *Pacific Linguistics*
*WPLUH* = *Working Papers in Linguistics, University of Hawaii*


ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DeBRUM

1976 *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.

ANDERSEN, Henning

1973 Abductive and deductive change. *Language* 49:765-793.

BENDER, Byron W.

1967 Some Micronesian word lists. Mimeo. University of Hawaii.

1968 Marshallese phonology. *OL* 7/1:16-35.

1969 *Spoken Marshallese*. Honolulu: University of Hawaii Press.

1969 Vowel dissimilation in Marshallese. *WPLUH* 1/1:88-96.

1971 Micronesian languages. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 426-465. The Hague: Mouton.

1973 Parallelisms in the morphophonemics of several Micronesian languages. *WPLUH* 5/8:1-16. Reprinted in *OL* 12 (1973):455-477.

1975 Mapian: an extinct Micronesian language? Paper presented at the Second Annual Oceanic Linguistics Festival, Linguistic Society of Hawaii, Honolulu.

1975 Stress in Marshallese. MS. University of Hawaii.

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

1981 A fossilized article in Marshallese. In Jim Hollyman and Andrew Pawley, eds *Studies in Pacific languages & cultures: in honour of Bruce Biggs*, 209-228. Auckland: The Linguistic Society of New Zealand.

BENDER, Byron W. et al.

1982 Proto-Micronesian word-list. Computer printout. Department of Linguistics, University of Hawaii.




(back matter) Bender, B.W. editor. *Studies in Micronesian linguistics.*
**C-80**, xii + 527 pages. Pacific Linguistics, The Australian National University, 1984. DOI:*10.15144/PL-C80.back*

BENTON, Richard A.

1968 Numeral and attributive classifiers in Trukese. *OL* 7/2:104-146.

BERRY, Jack

1958 The making of alphabets. In *Proceedings of the VIII International Congress of Linguists, Oslo,* 752-764(770). Oslo: University Press.

BIGGS, Bruce

1965 Direct and indirect inheritance in Rotuman. In G.B. Milner and Eugénie A. Henderson, eds *Indo-Pacific linguistic studies,* part 1: *Historical linguistics,* 383-415. Amsterdam: North-Holland. (=*Lingua* 14:383-415.)

BIGGS, Bruce, D.S. WALSH, and Jocelyn WAQA

1970 *Proto-Polynesian reconstructions with English to Proto-Polynesian finder list.* Interim listing January 1970. *Working Papers in Linguistics,* Department of Anthropology, University of Auckland.

BINGHAM, Hiram

1908 *A Gilbertese-English dictionary.* Boston: American Board of Commissioners for Foreign Missions.

1922 *Outlines of a grammar of the Gilbert Islands language.* Rongorongo, Beru, Gilbert Islands: London Missionary Society. (Originally compiled in 1861 in MS form only.)

BLUST, Robert A.

1972 Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages: a preliminary list. *WPLUH* 4/1:1-43.

1978 *The Proto-Oceanic palatals.* Memoir No. 43. Wellington: The Polynesian Society.

n.d. Variation in retention rate among Austronesian languages. MS.

BRANDSSTETTER, Renward

1916 *An introduction to Indonesian linguistics.* Translated by C.O. Blagden. London: Royal Asiatic Society.

BRESNAN, Joan

1977 Variables in the theory of transformations. In Peter W. Cullicover, Thomas Wasow, and Adrian Akmajian, eds *Formal syntax,* 157-196. New York: Academic Press.

BURLING, Robbins

1976 Poor spellers of the world, unite! *Learning* (March):76-78.

CAPELL, Arthur

1950 *A grammar of the language of Palau.* CIMA Report 6b. (Mimeo). Washington: Pacific Science Board.

n.d. (1951) *Grammar and vocabulary of the language of Sonsorol-Tobi.* CIMA Report 30. (Mimeo). Washington: Pacific Science Board. (Republished, 1969, as *Oceania Linguistic Monograph* 12.)

CARLSON, Clayton H.

1967 *Lessons in Palauan*. Revised edition. (Mimeo.) Honolulu: University of Hawaii Peace Corps Training Center (PALI).

1968 Palauan phonology. Master's thesis, University of Hawaii, Honolulu.

CARROLL, Vern

1966 Generative elicitation techniques in Polynesian lexicography. *OL* 5/2:59-70.

CARROLL, Vern and Tobias SOULIK

1973 *Nukuoro lexicon*. Honolulu: University Press of Hawaii.

CHOMSKY, Carol

1970 Reading, writing, and phonology. *Harvard Educational Review* 40/2:287-309.

CHOMSKY, Noam

1964 *Current issues in linguistic theory*. Janua Linguarum 38. The Hague: Mouton.

1965 *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

CHOMSKY, Noam and Morris HALLE

1968 *The sound pattern of English*. New York: Harper & Row.

CHOWNING, Ann

1973 Milke's 'New Guinea cluster': the evidence from northwest New Britain. *OL* 12:189-244.

CODRINGTON, Robert H.

1885 *The Melanesian languages*. Oxford: Clarendon Press.

CONSTANTINO, Ernesto

1971 *Ilokano dictionary*. PALI Language Tests. Honolulu: University of Hawaii Press.

COWELL, Reid

1951 *The structure of Gilbertese*. Beru, Gilbert Islands: Rongorongo Press.

DAVIS, Bonnie (Bonnie Davis ARAMAN)

1977 Who needs an alphabet? *WPLUH* 9/1:11-18.

DELATTRE, Pierre, Alvin M. LIBERMAN, Franklin S. COOPER, and Louis J. GERSTMAN

1952 An experimental study of the acoustic determinants of vowel colour: observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 8:195-210.

DEMPWOLFF, Otto

1934-38 *Vergleichende Lautlehre des austronesischen Wortschatzes*. 3 vols. Berlin: Reimer.

DOWNING, Bruce T.

n.d. English relativization. Mimeo. Department of Linguistics, University of Minnesota.

DYEN, Isidore

1949 On the history of the Trukese vowels. *LG* 25:429-436.

1965 *A sketch of Trukese grammar.* Essay 4. New Haven: American Oriental Society.

ELBERT, Samuel H.

1947 Ulithi-English and English-Ulithi word list. Mimeo. United States Naval Military Government.

1972 *Puluwat dictionary. PL*, C-24.

1974 *Puluwat grammar. PL*, B-29.

ELBERT, Samuel H. and Mary Kawena PUKUI

1979 *Hawaiian grammar.* Honolulu: University Press of Hawaii.

FISCHER, John L.

1965 The stylistic significance of consonantal sandhi in Trukese and Ponapean. *American Anthropologist* 67/6(1):1495-1502.

FLORA, Marie Jo-Ann

1969 Analysis of the segmental phonemes of Palauan. *Languages and Linguistics* 4:1-30.

1974 Palauan phonology and morphology. Doctoral dissertation, University of California at San Diego.

FOLEY, James

1972 Rule precursors and phonological change by meta-rule. In Robert P. Stockwell and Ronald K.S. Macauley, eds *Linguistic change and generative theory*, 96-100. Bloomington: Indiana University Press.

FRITZ, Georg

1911 *Die Zentralkarolinische Sprache: Grammatik, Übungen und Wörterbuch der Mundart der westlich von Truk liegenden Atolle, insbesondere der Saipan-Karoliner.* Berlin: G. Reimer.

GARVIN, Paul L.

n.d.
(1951) A linguistic study of Ponape. MS. CIMA, 1947-49.

1962 Ponapean morphophonemics. *Phonetica* 8/1-3:115-127.

GERAGHTY, Paul A.

1979 Topics in Fijian language history. Doctoral dissertation, University of Hawaii, Honolulu. Published as

1983 *The History of the Fijian languages. OL* Special Publication 19.

GIRSCHNER, Regimental Physician Dr

1906 Grammatik der Ponapesprache. *Mitteilungen des Seminars für Orientalische Sprachen:* I Abteilung, 73-126. Berlin. (Translated by S. Tully, 1975, as 'Ponapean grammar'. Mimeo.)

GOODENOUGH, Ward H.

1963 The long or double consonants of Trukese. *The proceedings of the Ninth Pacific Science Congress, 1957*, vol.3:77-86.

GOODENOUGH, Ward H, and Hiroshi SUGITA

1972 Final report of the conference to decide an orthography for the Truk language. Mimeo. University of Hawaii.

1980 *Trukese-English dictionary.* Memoir 141 of the American Philosophical Society. Philadelphia: American Philosophical Society.

GORBET, Larry

1974 Relativization and complementation in Digueño. Doctoral dissertation, University of California at San Diego.

GRACE, George W.

1969 A Proto-Oceanic finderlist. *WPLUH* 1/2:39-84.

n.d. On the history of the Trukese sound system. MS. University of Hawaii.

n.d. Some Trukese cognates (from Dyen 1949). Mimeo. University of Hawaii.

GRACE, George W., James A. THARP, Peter C. LINCOLN, and Janet BLACK

1979 Proto-Oceanic cognate sets. Computer printout. 202 pp. Honolulu: Department of Linguistics, University of Hawaii.

1980 Proto-Oceanic finderlist. Computer printout. 88 pp. [An earlier version appeared as Grace 1969.]

GULICK, Luther H.

1880 A vocabulary of the Ponape dialect: Ponape-English and English-Ponape. *Journal of the American Oriental Society* 10:1-109.

GUNDEL, Jeanette K.

1974 The role of topic and comment in linguistic theory. Doctoral dissertation, University of Texas at Austin. Reproduced by Indiana University Linguistics Club.

GUNDEL, Jeanette K. and Roderick A. JACOBS

1980 Why "seem" and "be" aren't what they seem to be. *Centrepoint* 4/2:115-123.

HALLE, Morris

1969 Some thoughts on spelling. In Kenneth S. Goodman and James T. Fleming, eds *Psycholinguistics and the teaching of reading,* 17-24. Newark: International Reading Association.

HAN, Mieko Shimizu

1962 *Japanese phonology: an analysis based upon sound spectrograms.* Tokyo: Kenkyusha.

HARRISON, Sheldon P.

1973 Reduplication in Micronesian languages. *OL* 12:407-454.

1977 Some problems in the history of Mokilese morpho-syntax. Doctoral dissertation, Department of Linguistics, University of Hawaii, Honolulu.

1978 Transitive marking in Micronesian languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings,* 1067-1127.

1983 Segmental quantity in Mokilese: a synchronic and diachronic study. See pp.375-401 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics*. *PL*, C-80.

in preparation *A grammar of Gilbertese*.

HARRISON, Sheldon P. and Salich W. ALBERT

1976 *Mokilese reference grammar*. Honolulu: University Press of Hawaii.

1977 *Mokilese-English dictionary*. Honolulu: University Press of Hawaii.

HARRISON, Sheldon P. and Frederick H. JACKSON

in preparation The evidence for a Micronesian grouping.

HOOPER, Joan B.

1973 Aspects of natural generative phonology. Doctoral dissertation, University of California at Los Angeles. Reproduced by Indiana University Linguistics Club, Bloomington, Indiana.

1976 *An introduction to natural generative phonology*. New York: Academic Press.

HOUSEHOLDER, Fred W.

1964 Three dreams of Modern Greek phonology. In Robert Austerlitz, ed. *Papers in memory of George C. Pappageotes*, 17-27. Supplement to *Word* 20/3. Special Publication No. 5. New York: Linguistic Circle of New York.

1967 Distinctive features and phonetic features. In *To honor Roman Jacobson: essays on the occasion of his seventieth birthday*, vol.2: 941-944. The Hague: Mouton.

HSU, Robert W.

1960 Palauan phonemics. Master's thesis, Georgetown University, Washington, D.C.

HYMAN, Larry M.

1975 *Phonology: theory and analysis*. New York: Holt, Rinehart and Winston.

JACKSON, Frederick H.

1976 Summary of decisions reached by Carolinian orthography committee. Mimeo.

1978 The synchronic status of the historical rule of final vowel deletion in Saipan Carolinian. Paper presented at the Historical Linguistics Festival of the Linguistic Society of Hawaii. Honolulu.

1983 Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. See pp.259-280 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics*. *PL*, C-80.

in preparation Aspects of the historical development of the Trukic languages of Micronesia. Doctoral dissertation, University of Hawaii, Honolulu.

JACKSON, Frederick H. et al.

forth-coming *Carolinian-English dictionary*. Honolulu: University Press of Hawaii.

JACOBS, Roderick A.

1973 Syntactic compression and semantic change. In Claudie Corum, T. Cedric Smith-Stark, and Ann Weiser, eds *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, April 13-15, 1973,* 232-241. CLS.

1975 Promotion and thematisation processes in English, or, How to get a head. In Robin E. Grossman, L. James San, and Timothy J. Vance, eds, *Papers from the Parasession on functionalism, April 17, 1975,* 223-231. CLS.

1975 *Syntactic change: a Cupan (Uto-Aztecan) case study*. Berkeley: University of California Press.

1976 A passive continuum in Austronesian. In Sandford B. Steever, Carol A. Walker, and Salikoko S. Mufwene, eds *Papers from the Parasession on diachronic syntax, April 22, 1976,* 118-125. CLS.

1976 Syntactic reconstruction and the comparative method: a Uto-Aztecan case study. In Mohammad Ali Jazayery, Edgar C. Polomé, and Werner Winter, eds *Linguistic and literary studies in honor of Archibald A. Hill,* vol.1: *General and theoretical linguistics,* 165-174. Lisse: Peter De Ridder Press.

1977 Syntactic categories and predictions in Gilbertese. Paper presented to the Linguistic Society of America Summer Meeting, Honolulu, Hawaii.

1980 Towards a semantic grammar of English. MS.

1983 Relatives, equatives, and information structuring. See pp.37-51 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics*. *PL*, C-80.

JENSEN, John Thayer

1977 *Yapese-English dictionary*. Honolulu: University Press of Hawaii.

1977 *Yapese reference grammar*. Honolulu: University Press of Hawaii.

JENSEN, John Thayer and Bruce BIGGS

n.d. (ca.1968) Three-letter abbreviations of Oceanic language names. Mimeo.

JOHNSON, Lawrence

1973 Dissimilation as a natural process in phonology. *Stanford Occasional Papers in Linguistics* 3:45-56.

JOSEPHS, Lewis S.

1975 *Palauan reference grammar*. Honolulu: University Press of Hawaii.

KAJITA, Masaru

1968 *A generative-transformational study of semi-auxiliaries in present-day American English*. Tokyo: Sanseido.

KARTTUNEN, Lauri

1970 On the semantics of complement sentences. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society, April 16-18, 1970,* 328-339. CLS.

KLIMA, Edward S.

1972 How alphabets might reflect language. In James F. Kavanagh and Ignatius G. Mattingly, eds *Language by ear and by eye: the relationships between speech and reading,* 57-80. Cambridge, Mass.: MIT Press.

KROHN, Robert

1971 On the sequencing of tautosegmental features. *WPLUH* 3/5:17-25.

KUROIWA, Kenneth

1973 Report on the Yap Outer Islands Orthography Committee Meeting. Mimeo. University of Hawaii.

1975 BO and LA: factive and non-factive complementation in Ulithian. MS. University of Hawaii.

LEE, Gregory

1976 Rule limitation in phonology. MS.

LEE, Kee-dong

1974 Kusaiean verbal derivation rules. Doctoral dissertation, University of Hawaii, Honolulu.

1975 *Kusaiean reference grammar.* Honolulu: University Press of Hawaii.

1976 *Kusaiean-English dictionary.* Honolulu: University Press of Hawaii.

LEE, Kee-dong and Judith W. WANG

1983 Kosraean reflexes of Proto-Oceanic phonemes. See pp.403-442 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics. PL,* C-80.

LESTER, Mark

1974 Bilingual education in the United States, the Pacific and Southeast Asia. *Topics in Culture Learning* 2:137-146. (Reprinted in R.W. Brislin, ed. *Culture learning,* 79-88. Honolulu: East-West Center.)

LEVY, Richard

1979 Languages of the southeast Solomon Islands and the reconstruction of Proto-Eastern-Oceanic. In Paz Buenaventura Naylor, ed. *Austronesian Studies: papers from the Second Eastern Conference on Austronesian Languages,* 213-225. Michigan Papers on South and Southeast Asia, No. 15. Ann Arbor: University of Michigan.

LIEBER, Michael D. and Kalio H. DIKEPA

1974 *Kapingamarangi lexicon.* Honolulu: University Press of Hawaii.

LYONS, John

1977 *Semantics,* vol.2. Cambridge: Cambridge University Press.

LYOVIN, Anatole

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

MARCK, Jeffrey C.

1977 A preliminary manual of nuclear Micronesian sound correspondences. Paper presented to the Austronesian Symposium, Linguistic Institute of the Linguistic Society of America, University of Hawaii, Summer 1977.

n.d. Thoughts on Proto-Trukic. MS.

McCAWLEY, James D.

1968 *The phonological component of a grammar of Japanese*. The Hague: Mouton.

McKAUGHAN, Howard P. and Batua A. MACARAYA

1967 *A Maranao dictionary*. Honolulu: University of Hawaii Press.

McMANUS, Edwin G., S.J.

1955 Word list and grammar notes - Palauan-English and English-Palauan. Mimeo. Koror, Palau. (See McManus 1977.)

1977 *Palauan-English dictionary*, ed. and expanded by Lewis S. Josephs, with the assistance of Masa-aki Emesiochel. Honolulu: University Press of Hawaii.

ODA, Sachiko

1977 The syntax of Pulo Annian: a nuclear Micronesian language. Doctoral dissertation, University of Hawaii, Honolulu.

PÄTZGOLD, Klaus

1968 *Die Palau-Sprache und ihre Stellung zu anderen indonesischen Sprachen*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 6. Berlin: Dietrich Reimer.

PAWLEY, Andrew K.

1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Readings in Oceanic culture history*, vol.3: 1-142. Pacific Anthropological Records 13. Honolulu: Bernice Pauahi Bishop Museum.

1977 On redefining 'Eastern Oceanic'. MS.

1979 Proto-Oceanic terms for people: a problem in semantic reconstruction. MS.

1979 *New evidence on the position of Rotuman. Working Papers in Anthropology, Linguistics, Maori Studies* No.56. Department of Anthropology, University of Auckland.

n.d. A Proto-Oceanic grammar. MS.

PIKE, Kenneth L.

1966 *Phonemics*. (10th printing.) Ann Arbor: University of Michigan Press.

QUACKENBUSH, Edward M.

1968 From Sonsorol to Truk: a dialect chain. Doctoral dissertation, University of Michigan, Ann Arbor.

n.d. Comparative Trukic word list. MS.

QUACKENBUSH, Hiroko C.

1970 Studies in the phonology of some Trukic dialects. Doctoral dissertation, University of Michigan, Ann Arbor.

RAY, Sidney H.

1926 *A comparative study of the Melanesian languages.* Cambridge: Cambridge University Press.

READ, Charles

1971 Pre-school children's knowledge of English phonology. *Harvard Educational Review* 41/1:1-34.

1973 Children's judgments of phonetic similarities in relation to English spelling. *Language Learning* 23/1:17-38.

1975 *Children's categorization of speech sounds in English.* Urbana: National Council of Teachers of English.

REHG, Kenneth L.

1973 On the history of Ponapean phonology. *WPLUH* 5/8:17-55. See also Rehg 1983, pp.281-316 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics. PL,* C-80.

1976 [Untitled]. Paper delivered at Linguistic Society of Hawaii panel discussion on optimal orthographies. May, 1976.

1978 Toward a theory of Ponapean prosody. Paper presented at the Linguistic Society of Hawaii Fourth Annual Oceanic Linguistics Festival.

1979 The origins of 'compensatory lengthening' rules in Micronesian languages. (Paper presented at the Linguistic Society of Hawaii Diachronic Linguistics Festival, November 1979.) See also Rehg 1983, pp.53-59 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics. PL,* C-80.

1981 *Ponapean reference grammar.* Honolulu: University Press of Hawaii.

1983 On the history of Ponapean phonology. See pp.281-316 in this volume, In Byron W. Bender, ed. *Studies in Micronesian linguistics. PL.* C-80.

1983 The origins of 'compensatory lengthening' rules in Micronesian languages, pp.53-59 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics. PL,* C-80.

REHG, Kenneth L. and Damian G. SOHL

1979 *Ponapean-English dictionary.* Honolulu: University Press of Hawaii.

REID, Lawrence A.

1976 *Bontok-English dictionary. PL,* C-36.

RIESENBERG, Saul H. and Shigeru KANESHIRO

1960 A Caroline Islands script. *Bureau of American Ethnology Bulletin* 173:269-333. Washington, D.C.: Smithsonian Institute.

ROSS, Malcolm

1977 Relationships of the Austronesian languages of the Sepik and western Madang coast of New Guinea. Mimeo.

1981 Proto-Oceanic from the top down. MS. University of Papua New Guinea.

SABATIER, Ernest, M.S.C.

1971 *Gilbertese-English dictionary.* (Translated by Sr Mary Oliva, F.N.D.S.C., of Tarawa, Gilbert Islands, Sacred Heart Mission.) Sydney: South Pacific Commission Publications Bureau.

SCHÜTZ, Albert J. and Tevita NAWADRA

1972 A refutation of the notion "passive" in Fijian. *OL* 11/2:88-109.

SJOBERG, Andrée G.

1966 Socio-cultural and linguistic factors in the development of writing systems for preliterate peoples. In William Bright, ed. *Sociolinguistics,* 260-276. The Hague: Mouton.

SMITH, Alfred G.

1951 Gamwoelhaelhi ishilh Weleeya (Guide to Woleai spelling). Mimeo. Department of Education, Office of the High Commissioner of the Trust Territory of the Pacific Islands.

SMITH, Frank

1973 *Psycholinguistics and reading.* New York: Holt, Rinehart and Winston.

SOHN, Ho-min

1971 *a*-raising in Woleaian. *WPLUH* 3/8:15-35.

1973 Relative clause formation in Micronesian languages. *OL* 12:353-392.

1975 *Woleaian reference grammar.* Honolulu: University Press of Hawaii.

SOHN, Ho-min and Byron W. BENDER

1973 *A Ulithian grammar. PL,* C-27.

SOHN, Ho-min and Anthony F. TAWERILMANG

1976 *Woleaian-English dictionary.* Honolulu: University Press of Hawaii.

SOHN, Ho-min, Anthony TAWERILMANG, Isaac IANGAL, and Celestine YANGILMAU

1977 Consonant shifts and subgrouping in the Sonsorol-Ulithi-Woleai chain. *Language Sciences* 44:19-24.

SOMMERSTEIN, Alan H.

1977 *Modern phonology.* London: Arnold/ Baltimore: University Park Press.

SOUTH PACIFIC COMMISSION

1981 *Economic development statistical survey.* Noumea.

SPOEHR, Alexander

1954 *Saipan: the ethnology of a war-devastated island.* Chicago: Natural History Museum.

STAHLKE, Herbert F.W.

1976 Which that. *LG* 52/3:584-610.

STEINBERG, Danny

1971 Would an orthography based on Chomsky and Halle's underlying phonological representations be optimal? *WPLUH* 3/3:2-18.

1977 Learning *kanji* is easier than learning *kana*. *WPLUH* 9/1:19-28.

STOCKWELL, Robert P. and J. Donald BOWEN

1965 *The sounds of English and Spanish.* Chicago: University of Chicago Press.

SUGITA, Hiroshi

1973 Semitransitive verbs and object incorporation in Micronesian languages. *OL* 12:393-406.

n.d. Lagoon Trukese grammar. MS. University of Hawaii.

in preparation Trukese reference grammar.

TAWERILMANG, Anthony F. and Ho-min SOHN

1983 Proto-Oceanic reflexes in Woleaian. See pp.171-199 in this volume, Byron W. Bender, ed. *Studies in Micronesian linguistics. PL,* C-80.

THARP, James A.

1976 *Itawis dictionary with English-Itawis finderlist.* New Haven, Connecticut: HRAFlex books.

TOPPING, Donald M.

1973 *Chamorro reference grammar.* Honolulu: University Press of Hawaii.

TOPPING, Donald M., Pedro M. OGO, and Bernadita C. DUNGCA

1975 *Chamorro-English dictionary.* Honolulu: University Press of Hawaii.

TRIFONOVITCH, Gregory J.

1968 Notes on Saipanese Carolinian grammar. MS.

TRUSSEL, Stephen

1979 *Kiribati (Gilbertese) grammar handbook.* Peace Corps Language Handbook. Brattleboro, Vermont: The Experiment in International Living.

TRYON, Darrell T.

1976 *New Hebrides languages: an internal classification. PL,* C-50.

WALLESER, Salvator

1911 Grammatik der Palausprache. In R. Lange and A. Forke, eds *Ostasiatische Studien,* 121-231. Berlin: G. Reimer

1913 *Palau Wörterbuch, I: Palau-Deutsch, II: Deutsch-Palau, nebst einem Anhange mit einigen Sprachübungen für Anfänger.* Hong Kong: Typis Societatis Missionum ad Exteros.

WANG, Judith W.

1981 Combined Oceanic reconstructions with English finderlist (preliminary version). Computer printout. Department of Linguistics, University of Hawaii.

WEIR, Ruth H. and Richard VENEZKY

1968 English orthography - more reason than rhyme. In Kenneth S. Goodman, ed. *The psycholinguistic nature of the reading process,* 185-199. Detroit: Wayne State University Press.

WELLEY, Weldis and Elaine GOOD

1976 A short grammar of Pingilapese. MS.

WILSON, Helen Irene

1972 *The phonology and syntax of Palauan verb affixes.* (Doctoral dissertation, Department of Linguistics, University of Hawaii, Honolulu.) *WPLUH* 4/5.

WILSON, Walter Scott

n.d. Outline of Kusaiean history. MS. Pacific Collection, University of Hawaii library.

WURM, S.A. and Lois CARRINGTON, eds

1978 *Second International Conference on Austronesian Linguistics: proceedings. PL,* C-61.

ZWICKY, Arnold

1972 Notes on a phonological hierarchy in English. In Robert P. Stockwell and Ronald K.S. Macauley, eds *Linguistic change and generative theory,* 275-301. Bloomington: Indiana University Press.